VOLUME 105 PART B NUMBER 19 JANUARY 1958

The Proceedings

OF

THE INSTITUTION OF ELECTRICAL ENGINEERS

FOUNDED 1871: INCORPORATED BY ROYAL CHARTER 1921

PART B

RADIO AND ELECTRONIC ENGINEERING
(INCLUDING COMMUNICATION ENGINEERING)

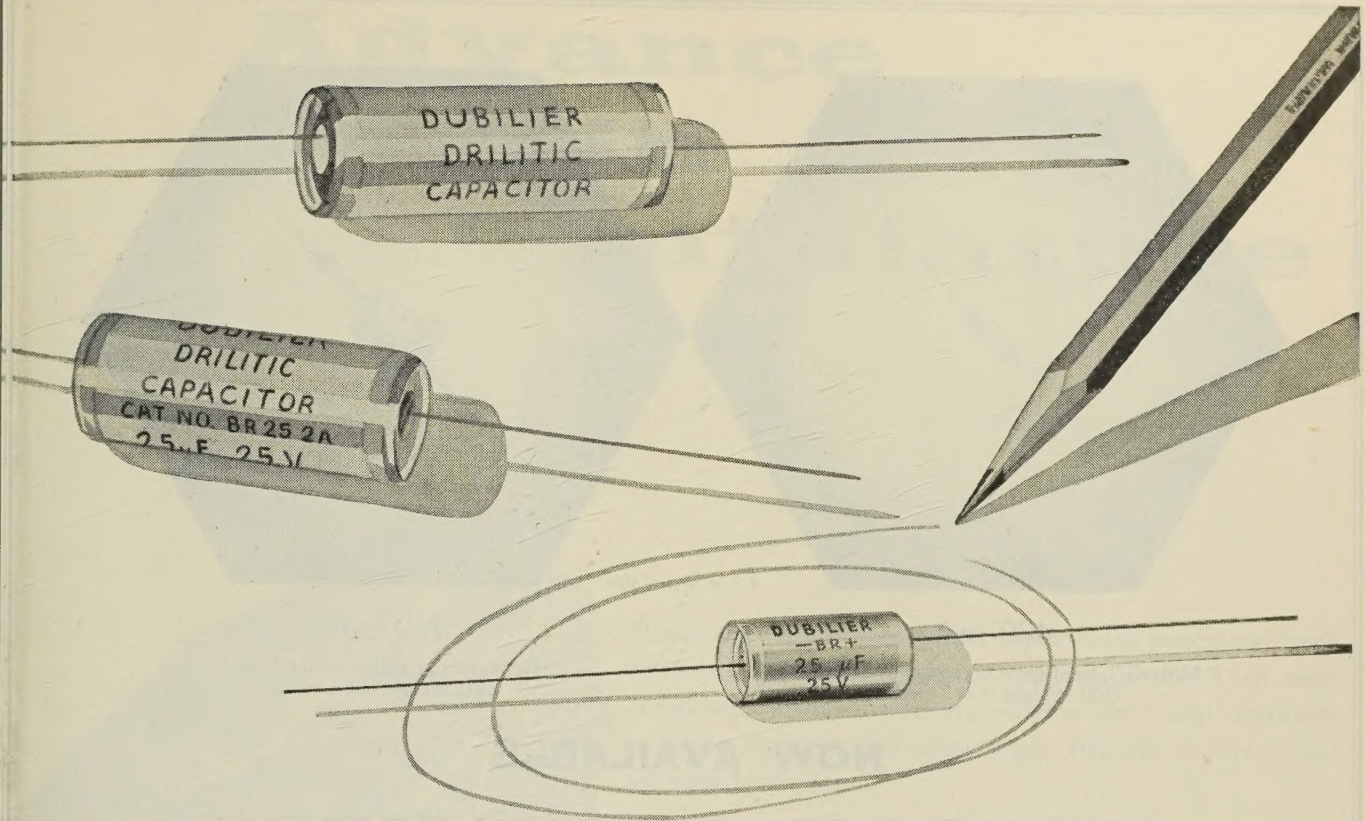
SAVOY PLACE · LONDON W.C.2

Price Ten Shillings and Sixpence

THE INSTITUTION OF ELECTRICAL ENGINEERS

FOUNDED 1871 INCORPORATED BY ROYAL CHARTER 1921

PATRON: HER MAJESTY THE QUEEN



we're reducing...

the size of our electrolytic capacitors. The new type BR miniature electrolytic capacitor retains all the characteristics of this well-known range with the added advantage of reduced size.

In modern equipment where every square inch of space is vital, these new capacitors will meet the physical requirements without any degradation of the electrical specification.

Capacitance (μF)	D.C. Wkg. Voltage	Dia. (in.) including sleeve	Length (in.)
25	12	$\frac{7}{16}$	$\frac{3}{4}$
50	12	$\frac{7}{16}$	$1\frac{1}{16}$
10	25	$\frac{7}{16}$	$\frac{3}{4}$
25	25	$\frac{7}{16}$	$\frac{3}{4}$
5	50	$\frac{7}{16}$	$\frac{3}{4}$
10	50	$\frac{7}{16}$	$\frac{3}{4}$
4	100	$\frac{7}{16}$	$\frac{3}{4}$
8	100	$\frac{7}{16}$	$1\frac{1}{16}$
2	150	$\frac{7}{16}$	$\frac{3}{4}$
4	150	$\frac{7}{16}$	$\frac{3}{4}$

We shall be pleased to supply full information upon request.

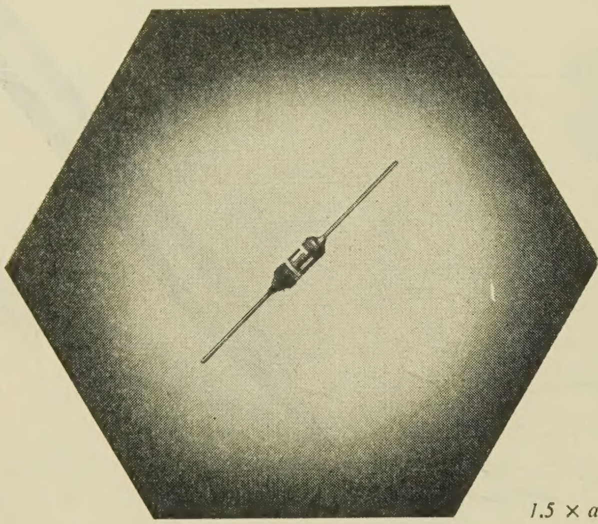
DUBILIER

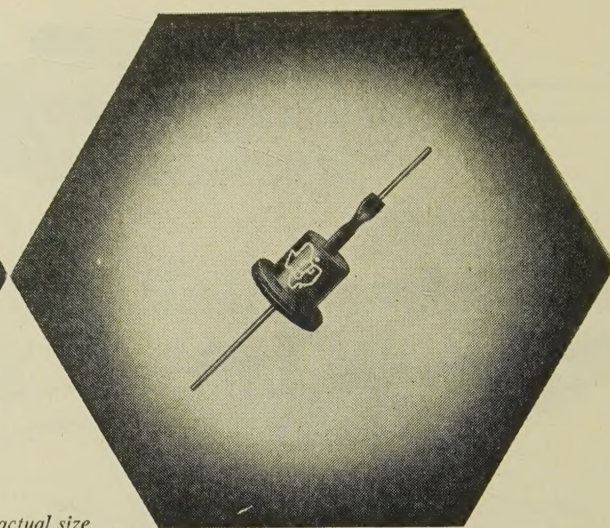
DUBILIER CONDENSER CO. (1925) LTD · DUCON WORKS · VICTORIA ROAD · NORTH ACTON · LONDON W.3

Telephone : ACOrn 2241

Telegrams : Hivoltcon Wesphone London

DN175B





1.5 × actual size

400 mA

"Moly-G", hard glass hermetically sealed case.
Unit Weight 0.195 Grms.

750 mA

Hermetically sealed metal case.
Unit Weight 1.6 Grms.

NOW AVAILABLE

SILICON RECTIFIERS

200—600 volts P.I.V.

These new 'Texas' Diffused Silicon Rectifiers provide you with :—

High current with high voltage
High Forward to reverse current ratio
Wide operating temperature range

TYPE NUMBERS	Peak Inverse Voltage:	200	300	400	500	600	volts
	Moly G:	1S111	1S112	1S113	1S114	1S115	
	Metal Case:	1S001	1S002	1S003	1S004	1S005	

		'MOLY G'	'METAL CASE'	
MAXIMUM RATINGS	Average rectified Forward Current at 25°C	400	750	mA
	Average rectified Forward Current at 150°C	150	250	mA
	Recurrent Peak Forward Current at 25°C	1.25	2.5	Amp.
	Surge Current 1 Sec. D.C. at +25°C to 150°C	3.0	15.0	Amp.
SPECIFICATIONS	Max. Reverse Current at P.I.V. 25°C	0.2	10.0	μA
	Max. Reverse Current at P.I.V. 100°C	15.0	300	μA
	Max. Voltage Drop at I_o=400 mA at 25°C	1.0	1.0	volt

Please write for Data Sheets of these Rectifiers and of our comprehensive range of Silicon Transistors.

TEXAS INSTRUMENTS LIMITED

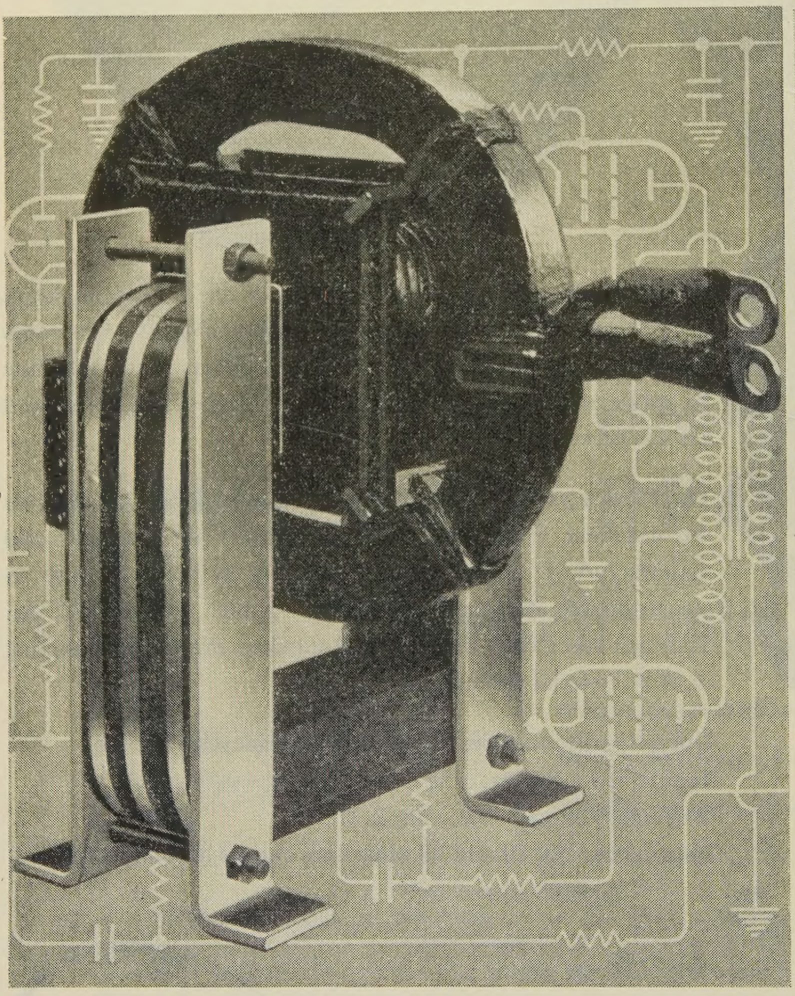
DALLAS ROAD · BEDFORD · TEL: BEDFORD 68051 · CABLES: TEXINLIM, BEDFORD

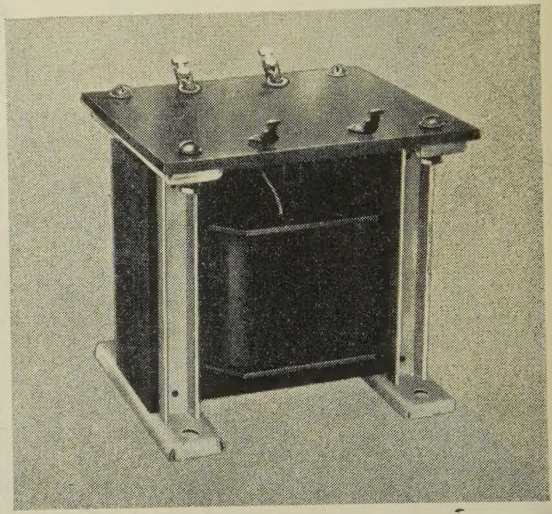
Extraordinary or Customary

Some Massicore transformers look very unusual for they are often the first of their kind anywhere and designed to meet unprecedented requirements.

Others have a more customary appearance, but *all* Massicore transformers have the same essential ingredient—conscientious craftsmanship. Massicore transformers are built to last a lifetime.

Equally important is the individual attention given to all enquiries and orders regardless of size; and we make a point of keeping our delivery promises.





The unusual Heater Transformer on the left was specially designed for use in circuits where a degree of isolation at radio frequencies is required. The input is standard 200/250 volts at 50 cycles, and the secondary is rated at 10 volts 200 amps. The special feature of this instrument is the extremely low capacity from the secondary to primary and core, this having been kept down to 24 p.f.

And above, more orthodox but made to the same unusually high standards, is a simple Switching Transformer. The primary is in series with the operating coil of a contactor and of an impedance high enough to prevent the latter closing until the secondary is short circuited. The secondary never develops more than 25 volts, can be earthed and operated by low voltage switch devices.

Corner for Contented Customers

"We are designing a new range of audio amplifiers for specialised work and knowing the high quality of your transformers wish to incorporate them in the design".

S.E., BOGNOR REGIS.

MASSICORE SAVAGE

SAVAGE TRANSFORMERS LIMITED. Devizes, Wiltshire. Tel: Devizes 932

TP/58

MARCONI TELEVISION RECORDING EQUIPMENT

TYPE BD679 RECORDING CHANNEL

16mm film presents a most economical and flexible means of recording television pictures since film costs are low, developing and printing techniques are advanced and a wide range of fine grain film stocks are readily available. Projection, editing, viewing, dubbing and handling are all easily carried out with standard equipment. The Fast Pull-down technique presents great advantages over other systems of recording on film but the mechanical difficulties of moving the film in the short period of frame blanking have hitherto prevented the employment of the technique. These problems have now been successfully overcome by Marconi's.

Features

Exceptionally high picture quality.

Specially developed gearbox enables Fast Pull-down technique to be employed. F.P.D. Mechanism has given over 3,000 hours trouble-free operation.

Pull-down time adjustable, normally set at 2 milliseconds permitting recording of fully interlaced picture. Simple single-lens optical system avoids loss of contrast.

Sound can be recorded on optical track, magnetic stripe or separate synchronous magnetic track.

Conveniently placed input selector switches, monitor and level controls. Sound/Vision cueing device incorporated.

Recording can be made on positive or negative stock of a wide variety, either direct positive, direct negative or reversal. Magazines hold 2,400 ft. (2,000 ft. magnetic stripe) or film may be fed directly into a rapid processor.

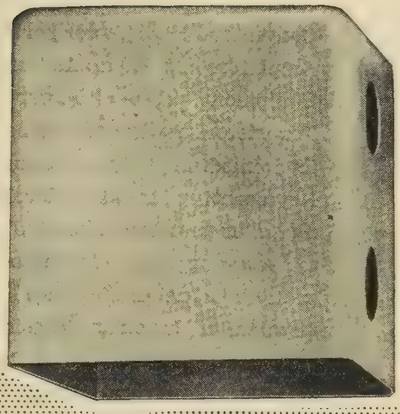
MARCONI

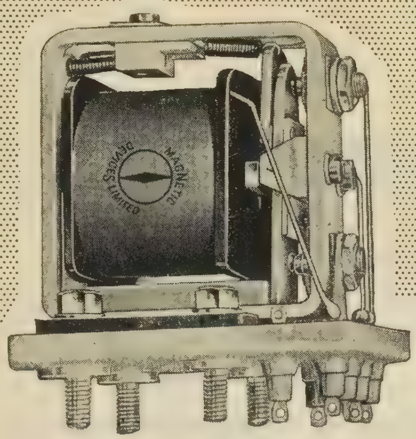
COMPLETE SOUND AND TELEVISION BROADCASTING SYSTEMS

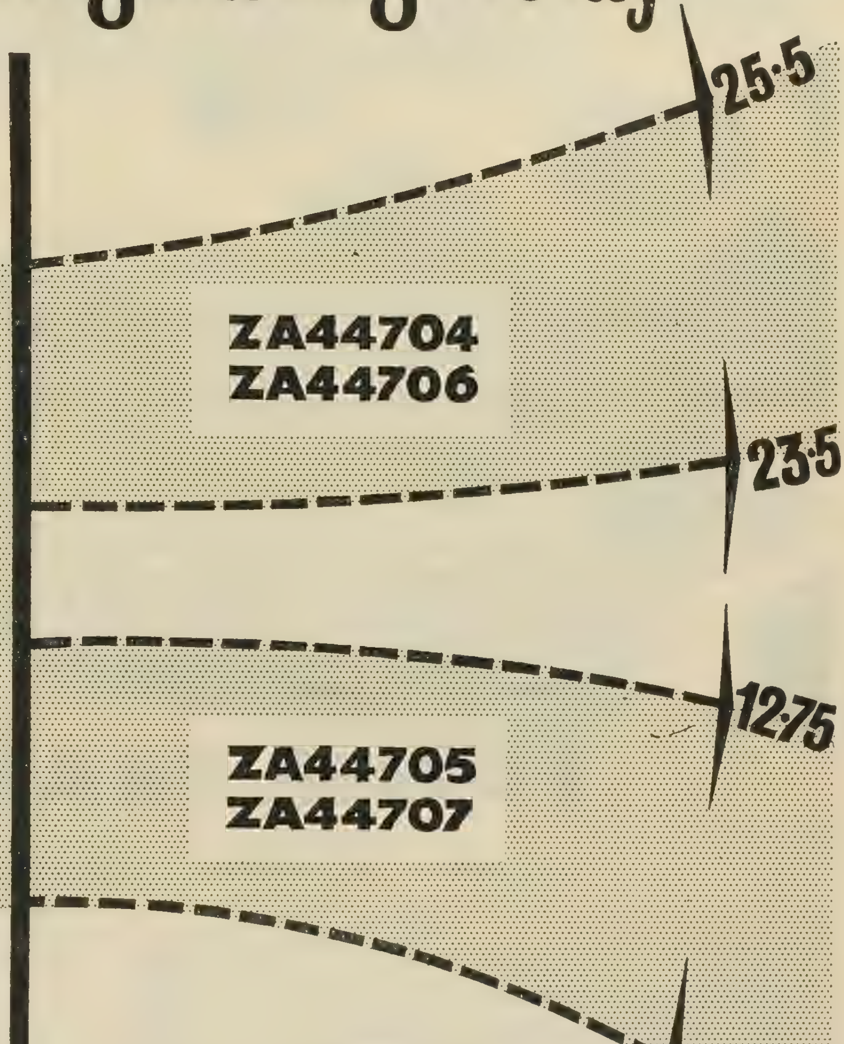
MARCONI'S WIRELESS TELEGRAPH COMPANY LIMITED, CHELMSFORD, ESSEX, ENGLAND

LB10

Voltage Regulating Relay







Two forms of the relay are available, either fully hermetically sealed or enclosed and tropicalised but unsealed. The inter-service reference numbers are as follows:—

Unsealed

ZA 44706 25.5V. make 23.5V. break
ZA 44707 12.75V. make 11.75V. break

Sealed

ZA 44704 25.5V. make 23.5V. break
ZA 44705 12.75V. make 11.75V. break

The Voltage Regulating Relay was designed in co-operation with S.R.D.E., to reduce voltage variations in certain essential circuits of radio sets and has many other applications of a similar nature.

This is particularly necessary in the case of vehicle-borne equipment, with power supplies consisting of lead acid batteries and a small charging generator.

The armature is balanced to withstand vibration, and the complete relay has been subjected to severe vibration testing.

Magnetic shielding is achieved by the iron case, enabling the relay to be used within reasonable proximity of transformers, chokes, etc.

The Voltage Regulating Relay complies with the stringent Ministry of Supply specification No. 166/1, to operate within tolerance, over a temperature range of – 40°C. to +85°C.

MAGNETIC DEVICES LIMITED

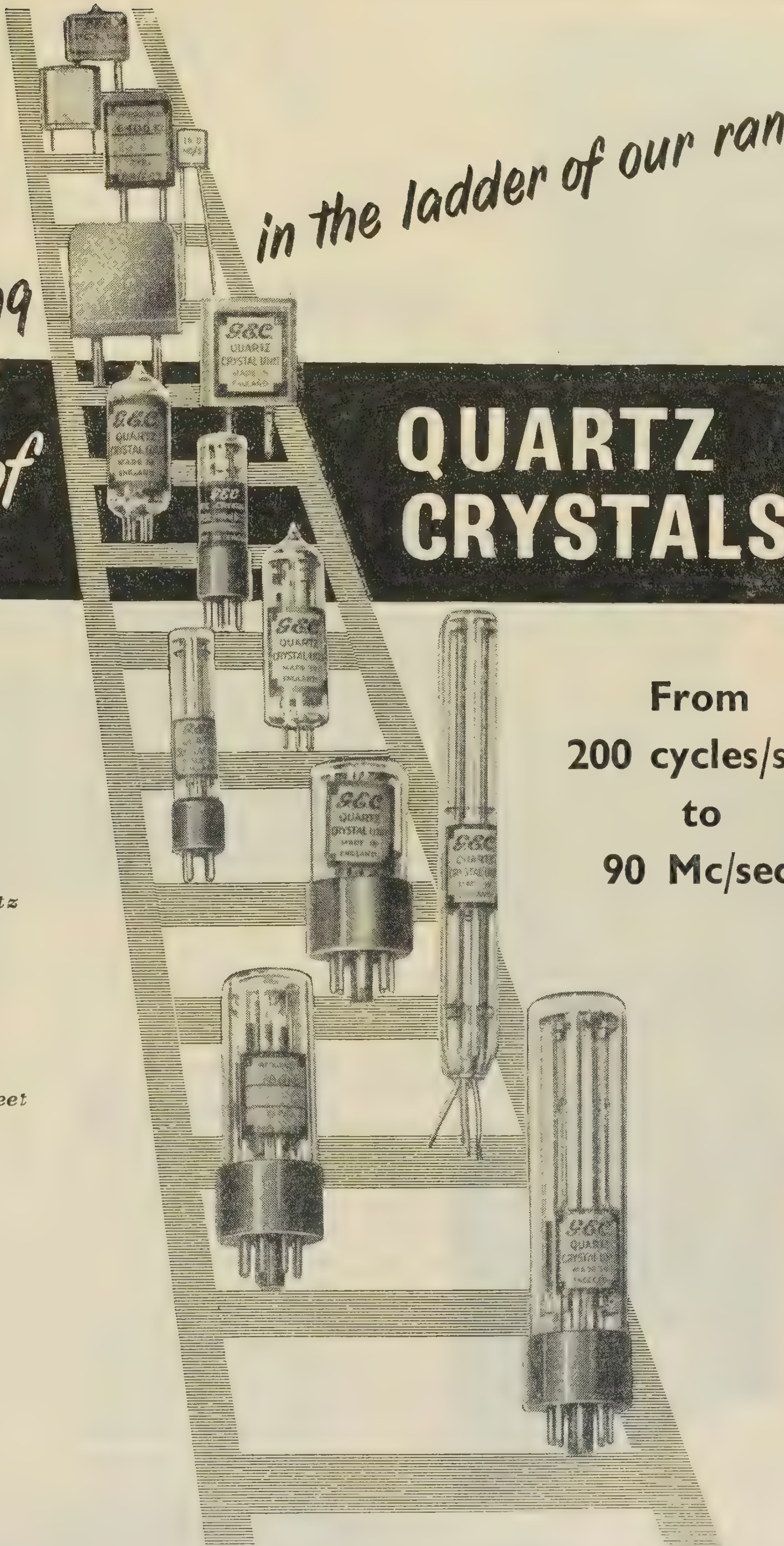
MAGNETIC DEVICES LIMITED

A.I.D. & A.R.B. approved

EXNING ROAD, NEWMARKET, SUFFOLK

Telephone: Newmarket 3181/2/3   Telegrams: Magnetic Newmarket

MD 20



no rungs missing in the ladder of our range

of **QUARTZ CRYSTALS**

G.E.C.

From 200 cycles/sec. to 90 Mc/sec.

For long term stability and unfailing activity, G.E.C. Quartz Crystal Units provide the basis for reliable communications systems.

A complete range of units to meet D.E.F.5271 and R.C.L.271 Inter-Services styles can be supplied.

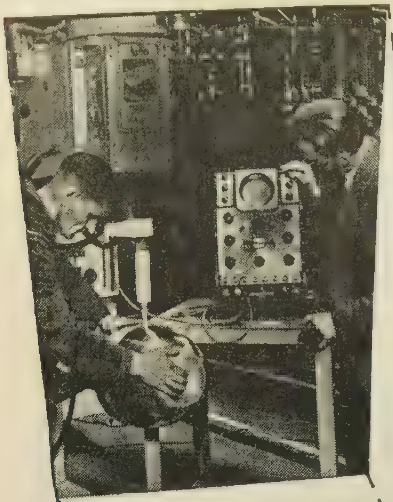
SALFORD ELECTRICAL INSTRUMENTS LIMITED

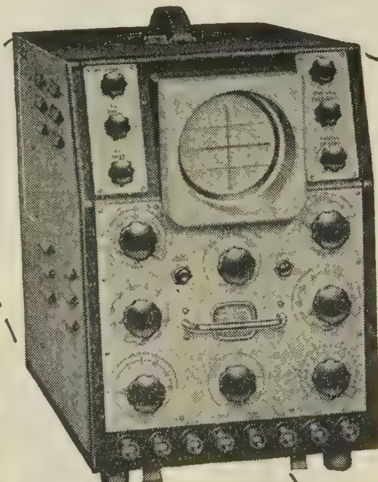
(COMPONENTS GROUP)

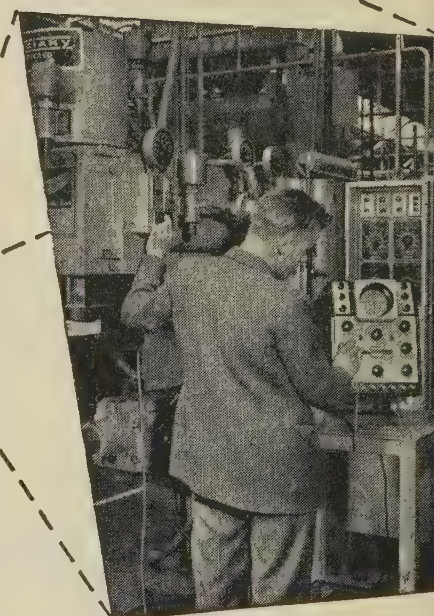
TIMES MILL · HEYWOOD · LANCASHIRE Tel: Heywood 6868

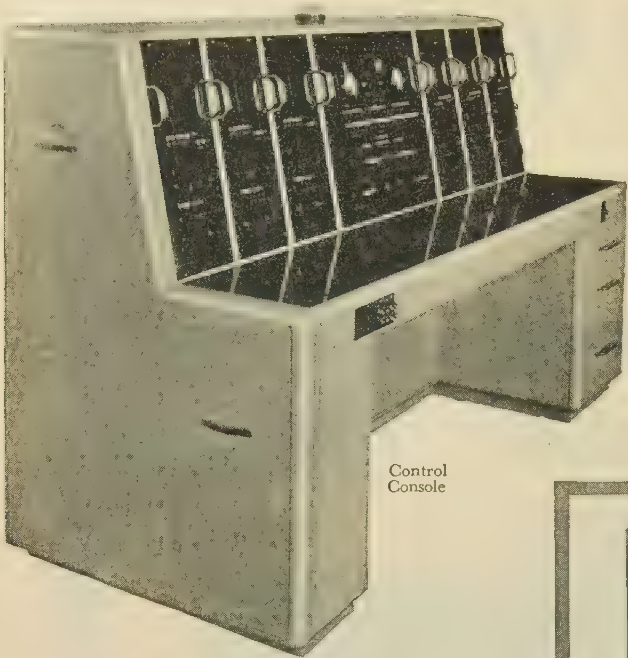
London Sales Office Tel: Temple Bar 4669

A SUBSIDIARY OF THE GENERAL ELECTRIC CO. LTD. OF ENGLAND





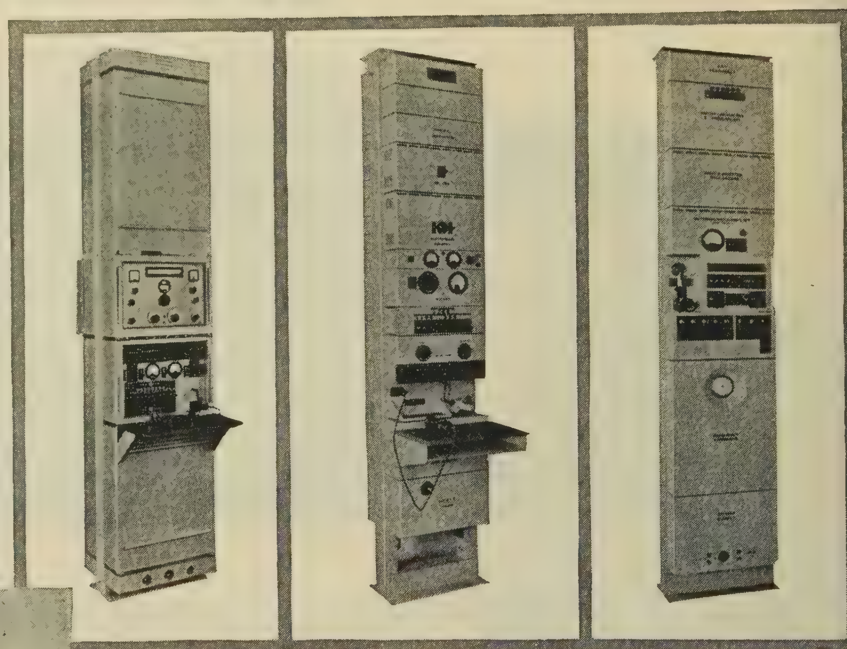




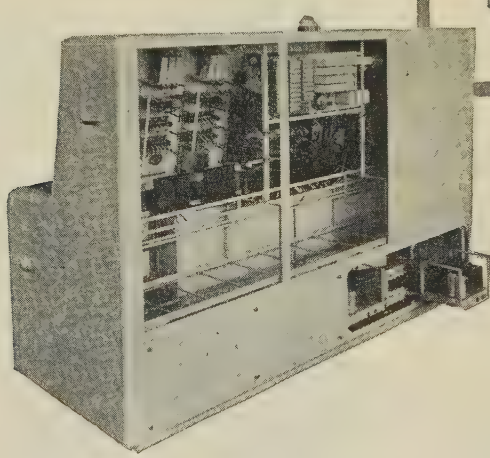
Control Console

simple or complex

* Electronically controlled 5 band privacy equipment.
* Channel shifting equipment for I.S.B. transmission.
* Centralised control, local or remote.
* Full monitoring and supervision facilities.



RTT 'F' RTT 'B' PRIVACY '5E'



Siemens Ediswan radio telephone terminal equipment provides efficient connection of a land telephone system to a radio transmitting and receiving station.

Privacy equipment of various types is obtainable, and allows a high degree of protection against unauthorised interception of speech on a radio telephone circuit.

There is a wide range of equipment available. Full information will gladly be sent on request.

extending → the frontiers of telecommunications

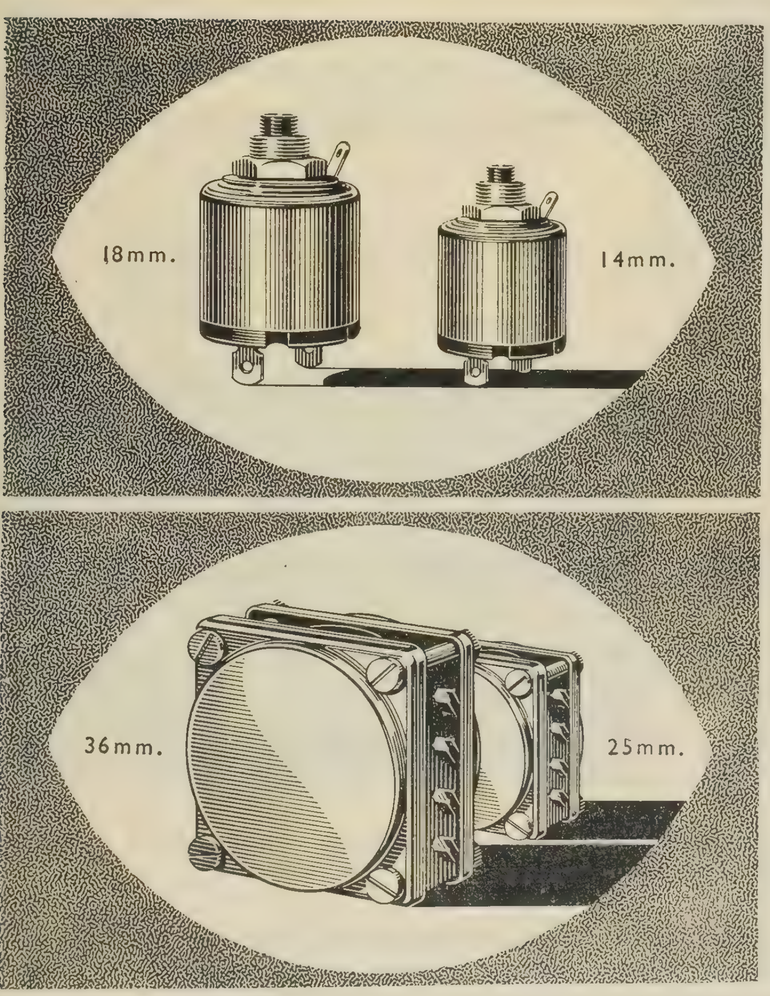
SIEMENS EDISWAN

SIEMENS EDISON SWAN LTD. *An A.E.I. Company*
Telecommunications Transmission Division, Woolwich, London, S.E.18. Telephone: Woolwich 202
Telegrams and Cables: Sieswan Souphone London.

TAB2

MULLARD HIGH EFFICIENCY POT CORES

have these outstanding features ....



* Pot core design facilitating rapid assembly
* Small size
* High value inductance
* Low losses resulting in high Q values
* Very fine setting accuracies
* Operative over a wide frequency range
* Controlled temperature coefficient

Wherever high quality pot cores are required, there will be a Mullard type available to meet the specification, furthermore, they can be supplied wound to customers individual requirements.

Write now for full details of the comprehensive range currently available.

Mullard

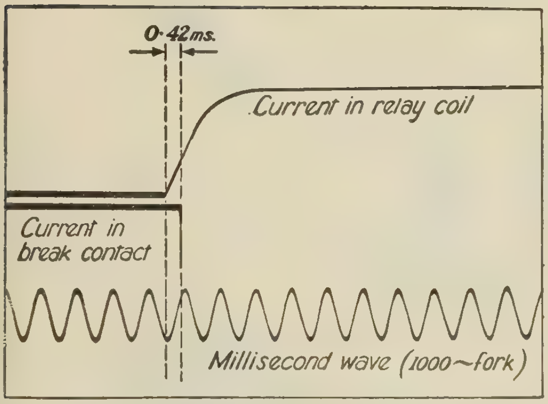
'Ticonal' permanent magnets
Magnadur ceramic magnets
Ferroxcube magnetic cores

MULLARD LTD., COMPONENT DIVISION, MULLARD HOUSE, TORRINGTON PLACE, W.C.I

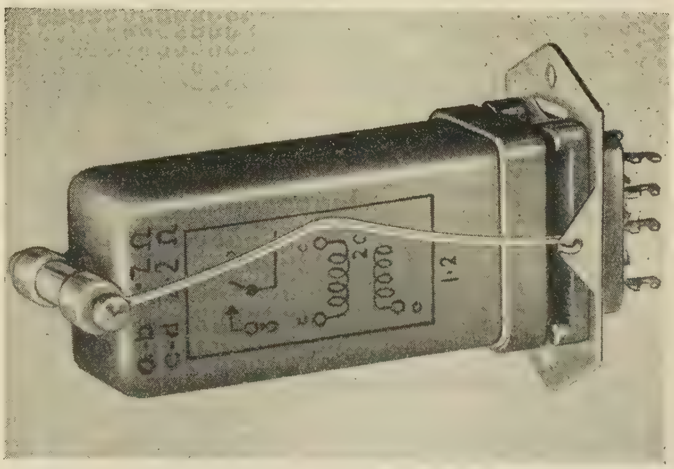
MC 255A

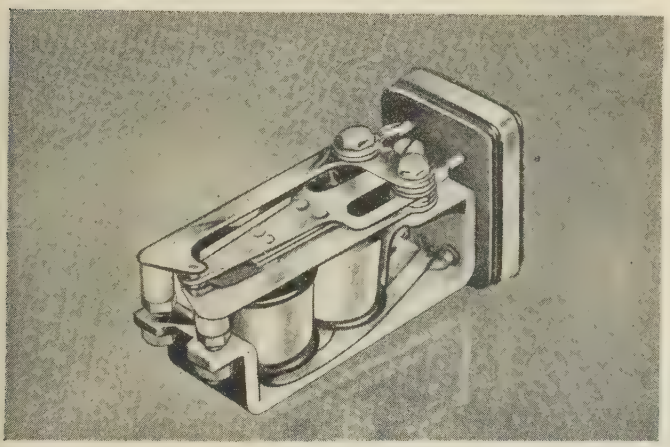
SIEMENS EDISWAN

Miniature
high speed relay



Oscillogram of the operating-lag of the break contact





This small relay has earned for itself a world-wide reputation on account of its very rapid operation, great reliability and insensitivity to external mechanical and electrical disturbances. It was originally designed to meet very severe conditions called for by the Services; it is now in use in vast numbers for all manner of applications and varying conditions.

Inexpensive and available for early delivery.

Can be supplied with plug-in base if required to form a readily interchangeable plug-in unit.

Hermetically sealed, unaffected by dirt or immersion in water and immune to wide or rapid changes of temperature or air pressure.

We shall be pleased to give full details if you telephone (Woolwich 2020, Extension 621) or write to:

SIEMENS EDISON SWAN LTD *An A.E.I. Company*

WOOLWICH LONDON SE18

TRANSISTOR OC16

Dissipation

14 watts at 45°C *

INDUSTRIAL APPLICATIONS

This high output germanium junction P.N.P. power transistor is being used to great advantage in the following applications:

Class A and B audio Amplification

Aircraft Passenger Address Systems

D.C. Convertors

Power Amplifiers for Servo Motors

Power for Relay Operation.

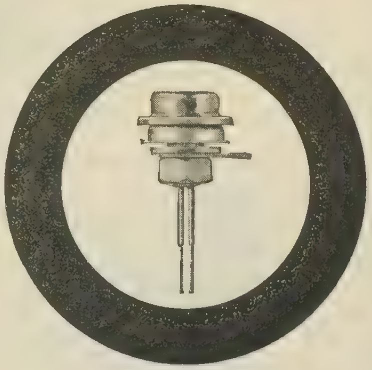
FEATURES

Peak current 3 amps.

Maximum mean current 1.5 amps with good current linearity.

Maximum continuous junction temperature 75°C. Up to 90°C for intermittent operation (total duration 200 hours max.).

A pair of OC16's operating in Class B push-pull will provide an output in excess of 10 watts into loudspeakers with 12 volts "H.T."



Half actual size

DATA

Characteristics (measured at 25°C ambient)

Collector reverse leakage current, grounded base .... <100μA

Current amplification
- at Ic = 30mA 40
- at Ic = 300mA 35
- at Ic = 2A 22
- at Ic = 3A 16

*** POWER DISSIPATION**

14 watts dissipation at an ambient temperature of 45°C can be allowed when the OC16 is bolted to a large but practical heat sink. The accompanying table shows allowable dissipations with different heat sinks. In all cases the transistor is mounted directly on to a plate with a thin tin-plated lead washer.
When electrical insulating mica washers are used, the dissipation is somewhat reduced. Full data is available.

AMBIENT TEMP. (TO GIVE 75°C JUNCTION TEMP.)	DISSIPATION WITH INFINITE HEAT SINK, e.g. WATER COOLED	DISSIPATION WITH LARGE METAL PLATE 1°C PER W THERMAL RESISTANCE	DISSIPATION WITH 1 sq. ft. METAL PLATE 2°C PER W THERMAL RESISTANCE	DISSIPATION WITH SMALL METAL PLATE (4 in. x 7 in.) 5.5°C PER W THERMAL RESISTANCE
25°C	45W†	24W	16W	7.6W
35°C	36W†	19W	13W	6.1W
45°C	27W†	14W	9.7W	4.6W
55°C	18W	9.5W	6.5W	3.0W

†Continuous dissipation values in excess of 24W are at present prohibited by the voltage and current ratings given in the full data.

MULLARD LIMITED
MULLARD HOUSE
TORRINGTON PLACE
LONDON W.C.1

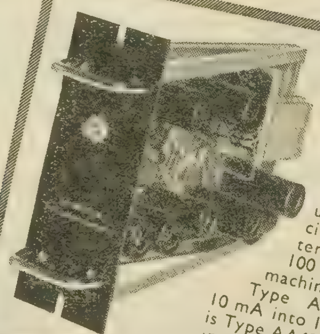
Mullard

COMMUNICATIONS AND INDUSTRIAL VALVE DEPARTMENT

MVT 329a

THE SOLARTRON COMPUTER SERVICE

YOU BUILD YOUR COMPUTER WITH OUR STANDARD UNITS

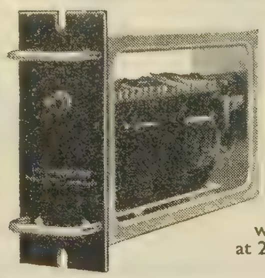


DRIFT-CORRECTED COMPUTER AMPLIFIER AA621

A high-performance D.C. amplifier of completely new design using the latest printed circuit techniques. Long term drift is less than 100 μV and output about machine datum is ± 100 V. Type AA621 will deliver 10 mA into 10KΩ; also available is Type AA622, delivering 50 mA into a 2KΩ load.

COMPUTER POWER SUPPLY AS755

Uses germanium junction rectifiers and provides 300 V at 0·5A with excellent output stability and low ripple content. Two units cater for the normal differential operation of 20 Computer Amplifiers.

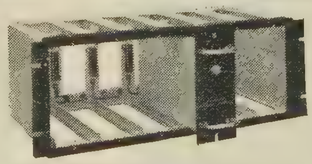


HEATER SUPPLY UNIT TS722

To ensure top stability and freedom from interaction between Computer Amplifiers each is supplied from an individual heater winding; output is 6·3V at 2A on each of 6 outlets.

VENTILATING PANEL TX724

A simple but effective Ventilating Panel which deflects hot air to the front and rear of assembled computer racking.



AMPLIFIER MOUNTING UNIT TX791

Houses five Computer Amplifiers and fits standard 19″ P.O. rack. Provided with mains and H.T. links, also patching and test socket panels.

Also available:

Linear Computing Unit TR729	*24V Relay Supply Unit*
Crossed Field Multiplier TA721	*100V Computer Reference Supply*
	Control Switching Panel

SOLARTRON

The Solartron Computer Service offers you a comprehensive range of precision "building bricks" for building your own analogue computer or simulator—or extending the machine you are already operating. All units in the range are capable of independent use if desired but have been designed around a standard 100V machine unit and use ± 300V H.T. supplies.

THE SOLARTRON ELECTRONIC GROUP LTD. THAMES DITTON. SURREY

Tel: Emberbrook 5522 **Cables: Solartron, Thames Ditton**

THE PROCEEDINGS OF
THE INSTITUTION OF ELECTRICAL ENGINEERS

EDITED UNDER THE SUPERINTENDENCE OF W. K. BRASHER, C.B.E., M.A., M.I.E.E., SECRETARY

VOL. 105. PART B. NO. 19. JANUARY 1958 ©

The Institution of Electrical Engineers
Paper No. 2436
Oct. 1957
©

INAUGURAL ADDRESS

By T. E. GOLDUP, C.B.E., President.

BACKGROUND, FOREGROUND AND HORIZON

The Radio Valve Industry in Prospect and Retrospect

(*Address delivered before* THE INSTITUTION *3rd October*, 1957.)

To be elected President is the highest honour that can be onferred on a member of this great Institution, and it is an vent of extreme importance to those who are chosen by their llow-members to serve in this capacity.

I am deeply appreciative of this honour, which I regard as a ost generous compliment, not only to me personally, but also the light electrical engineering branch of the profession, and deed to the radio industry which I have served for so long. uring my term of office the responsibilities I assume as your esident will be constantly in my thoughts as I strive to uphold e high standard of service which has characterized the work those distinguished electrical engineers who have occupied e Presidential Chair in past years.

There is not one of us, I am sure, who does not look with pride 1871 as the most significant year in the history of electrical gineering, for in that year The Institution came into being nder its original title of 'The Society of Telegraph Engineers'. recalling the state of science at that time, it is interesting to ote that it was in this same year that James Clerk Maxwell was ppointed the first Professor of Experimental Physics at Cam-idge, where he propounded his theories of the electromagnetic ature of radiant energy, and founded the Cavendish Laboratory. f the eight original founders of The Institution (five of whom cidentally were members of the Armed Forces), together with e sixty-six original members as recorded in the minutes of ay, 1871, none could possibly have foreseen the ultimate nportance of The Institution as a national body, and certainly one could have contemplated the possibility that in less than century the members, numbering more than 40 000, would ve elected a President whose working life had been spent in a ld described to-day as light electrical engineering and on a ly specialized subject, the threshold of which had not then en crossed. Such, however, is the scale of the time-base ncerned with the development of thermionic valves, the lden Jubilee of which was celebrated in this lecture theatre in 54. Our founders certainly could not have imagined the pact of Fleming's or de Forest's work on the pattern of our e, to-day, where thermionic valves and a whole family of related devices have become an all-important factor in every branch of engineering and science—a factor which more than any other has not only shaped but determined the progress in electrical engineering as we know it to-day, and has in consequence profoundly influenced almost every other form of human endeavour.

It is with these thoughts that I embark upon the task of delivering this Address and in so doing I would first like to dwell for a moment on the significance of The Institution in our national life. Our Royal Charter requires us as an Institution 'to promote the general advancement of Electrical Science and Engineering and their applications, and to facilitate the exchange of information and ideas on those subjects amongst members of The Institution'.

The whole machinery of The Institution is devoted to this end, an important part being the work of the Local Centres and Sub-Centres, and the Oversea Branches and Joint Groups, constituting 64% of the total membership. The meetings and discussions which these Centres and Groups organize are devoted directly to the implementation of the terms of our Royal Charter.

We all appreciate the wisdom of our predecessors in providing these facilities, which have done so much to meet the needs of members in the provinces and oversea, as well as adding to the prestige of The Institution as a whole. But a good deal more is involved than the mere dissemination of information among our members and the promotion of electrical science and engineering, to say nothing of our responsibility as a body to ensure and maintain adequate professional standards. Important and necessary as all this is, in the ultimate it is the collective effort of individuals which matters, and each of us has a personal responsibility to ensure that his own individual contribution to the profession is adequate in all respects. That the total contribution is adequate is reflected in the pre-eminent position our Institution is proud to hold, but it is nevertheless the fact that this contribution comprises the strenuous efforts of only a relatively small number of members, and I would ask each of you to consider whether your individual contribution to The Institution and to

the development of the science in which you practise is as great as you feel it should be.

We must never forget that it is the responsibility of the coming generation of engineers to see that our present leadership is maintained. Their task in this respect will demand a recognition of the fact that these are days when the knowledge which is increasing rapidly in all fields of engineering and science must be matched with increased wisdom, for if these two fail to go hand in hand disaster will eventually overtake us. I appeal to the younger members of this Institution to mark well my words and ponder on the truth expressed in Proverbs viii, 11:

> For wisdom is better than rubies and all the things that may be desired are not to be compared to it.

Now I must turn from The Institution—the frame in which my Address is set—to the picture, the Address itself, and I shall sketch it on a broad canvas, my main theme being thermionic valves, a subject which has constituted my life's work, first for a few years in the sheltered atmosphere of an Admiralty research establishment and then for many years in the hard but exhilarating school of industry. I am anxious regarding the scope and length of my Address for I am faced with the difficult task of deciding the extent of the ground to be covered.

That the major part of my Address should have as its theme some aspect of thermionic valve development I was never in doubt, and it was my original intention to give a more or less complete review of valve developments. After much thought, however, it became increasingly obvious that to include the whole story would make this Address far too long. I then considered at some length those points of valve history which seemed to me most significant. But such is the nature of this field that I found I was merely repeating what had already been said and written elsewhere far more ably than I could attempt, and often in this very theatre. For these reasons and because I wish my Address to break new ground and perhaps leave you with some new lines of thought, I intend to confine what I have to say to a few aspects which I consider of interest and which are unlikely to be familiar to those of you who do not work in the valve field.

BACKGROUND

On the 16th November, 1954, The Institution celebrated the Jubilee of Ambrose Fleming's invention of the thermionic valve by a series of lectures and an exhibition. The proceedings of this memorable occasion, together with a catalogue of the exhibits, are recorded in an Institution publication entitled 'Thermionic Valves 1904–1954', which gives a good pictorial representation of the progress achieved since Fleming applied for his patent in 1904. This record embodies the whole family of thermionic valves, from simple diodes to the multi-electrode valves, and on to magnetrons, klystrons and travelling-wave tubes, all of which bear little resemblance to the fragile hand-made diode which Fleming used in his Gower Street laboratory for his early experiments.

With all new products an evolutionary streamlining process occurs in which new design ideas, new materials and new applications all combine to influence appearance, utility and efficiency to such an extent that the original design is barely recognizable. The thermionic valve is no exception, its present appearance bearing little resemblance to the electric lamp from which it is in many respects descended. Its first appearance in the form of a simple diode opened a field of progress in design, manufacture and application, the horizon of which is still a long way off.

Manufacture of valves in this country began during the years of the First World War, when a new industry developed, born of the vacuum techniques which were in common use in electric lamp manufacture. The somewhat limited output of both vacuum and gas-filled types, referred to in the jargon of those days as 'hard' and 'soft' valves, was used exclusively for the communication equipment then available to the Armed Forces. However, apart from some specialized applications in telephony, the valve output of the early 1920's was used for broadcast receivers, and it has been the subsequent development of sound broadcast services and later television which has provided the incentive for the development of the valve industry.

There is little point in commenting upon the numerous types of valves now available—the figure is nevertheless somewhat staggering, as is also the multitude of applications that are covered. From the early days of manufacture there has been a relentless and continuous attack on valve design and manufacturing problems in an endeavour to meet the seemingly endless demands of existing and new applications which have become characteristic of the advance in the ever-widening field of electronics. As a result we now have thermionic devices of every conceivable type, from the conventional receiving and transmitting valves to their more complex counterparts for use at the highest frequencies; also counting, storage, display and infra-red devices, and the transistors that have emerged from the study of solid-state physics.

The progress achieved is an example of combined operations on a number of technical fronts, involving the physicist, the chemist and the metallurgist, as well as the electrical and the mechanical engineer. By this consortium of effort, the design, manufacturing and application problems have been tackled and solved. In reviewing the highly satisfactory results that have been achieved it is, strangely enough, most difficult to highlight specific technological achievements; so far as manufacturing technology is concerned there has been an advance on the broad fronts of chemical and mechanical engineering, combined with a strict control of raw materials and processes. The introduction of new and more suitable materials and the rigid adherence to manufacturing schedules have also contributed to the development of new and improved thermionic devices.

FOREGROUND—PLANT AND PRODUCTION

The following is a brief description of one or two production processes, intended to give some idea of the materials from which valves are made, their various component parts and how they are assembled.

Taking as an example an indirectly heated h.f. pentode, the cathode consists of a flat nickel tube which is coated with a mixture of barium and strontium carbonates 80 microns thick. When the valve is heated during the pumping process these carbonates are converted to oxides, and final processing produces a small amount of free barium and strontium, without which the emitter would not function. In addition, the valve design must be such that the emissive coating is not damaged by the heat applied to the valve envelope and electrodes during sealing and exhausting.

The heater, which fits inside the cathode, is made of tungsten wire 70 microns in diameter, drawn to an accuracy of $\pm 0{\cdot}7$ micron. Its operating current is 300 mA and it works at 1200° C.

The molybdenum wire used for the grids has a tolerance of $\pm 1\frac{1}{2}\%$ on diameter. The grid-pitch tolerance is ± 1 micron, and this accuracy must remain after the grids have passed through a cleaning process where they are heated in an atmosphere of hydrogen at 900° C.

Finally, these electrodes and the anode are assembled between two mica discs, which must locate them to an accuracy of 20 microns; after assembly the distance between the electrodes is nowhere more than half a millimetre.

The production of glass parts is of great importance in valve nd cathode-ray-tube manufacture. The need for accurate mensions and uniform composition requires a detailed application of glass technology. Glass tubing, in one application, is sed for cathode-ray-tube necks, the tube being flared at one end fit the cathode-ray-tube cone. This flaring is done at a mperature of 800° C and at a rate of four per minute. In nother application, to valve envelopes, a tube is closed at one nd in a mould at a temperature of 1 050° C.

In another process the face of a television tube is joined to the one. The glass is some $\frac{3}{8}$ in thick to withstand atmospheric essure when the tube is evacuated, and in order to establish the aling temperature uniformly across this width, after the glass as first been heated by gas flames, a current is passed through , taking advantage of the property of glass whereby its electrical sistance is relatively low at high temperatures.

While a vacuum device is being exhausted the glass and ectrodes are heated so that any absorbed gases are driven off. or instance, a large cathode-ray tube passes through a heated nnel during a period of some two hours; in the case of receiving alves this process is carried out in a few minutes on much maller machinery of an entirely different design. This demonrates the range of techniques employed in a single process.

While many production processes are done mainly or wholly machine, a great deal of work is carried out by semi-skilled operators. Here the technique of operator training replaces to large extent the technique of production machinery design, nd correct selection and training methods are just as important accurate know-how and technology. Methods of operator aining are always being given considerable thought, but we ave still quite a lot to learn.

The foregoing will, I hope, serve as an example of the combined forts of mechanical, chemical and electrical engineering plus ome applied physics, metallurgy and glass technology, all o-ordinated in the interest of efficient mass production of ermionic valves and kindred devices.

I included this account to illustrate the part that plant design nd development has played in the manufacture of valves. It a subject which seldom appears in papers read before The stitution, and for very good reasons, but this in no way inimizes its importance.

Research and development have their own special appeal and e a distinctive and satisfying element in the achievement of sults which thrive on differences of approach to the solution problems. The same is equally true of plant design and evelopment, and, as a stage in the chain of events leading to a anufactured article, their importance is paramount, for they e prime factors in the economics of production. In this onnection it is as well to remember that only by economic and ficient production can industry continue to exist and provide e necessary finance for research and development, and indeed r all its other activities. Additionally, plant design determines ry largely the actual manufacturing processes to be used and what extent it is possible to employ modern technological lvances.

The design problems of radio valves and for that matter of y thermionic device are, as in the case of most products, timately linked with the manufacturing possibilities, which in rn are mainly dictated by the capabilities of the plant and its gners.

New techniques are constantly being introduced, a recent novation being the working of metals by electro-sparking volving an electric erosion effect. This enables holes of the ost complex shape to be cut in the hardest metals by unskilled bour. An electrode, which can be of brass or steel, is first ade to the shape and size of the hole required, and by passing an electric discharge through a liquid dielectric medium between this electrode and the workpiece, a hole is cut corresponding to the outline of the electrode. This hole, cut to very close limits, is extremely smooth, the workpiece produced being suitable for incorporation in, say, a punch tool without any further finishing. Flame-plating is another technique with great possibilities which, by the use of very high temperatures and pressure, enables tungsten carbide to be plated on a great variety of metals, including steel, cast iron, aluminium, copper and brass. This enables a part to be produced with an extremely hard surface, but retaining the metallurgical properties of the base metal. Aluminium oxide can also be plated on by this process, producing a surface highly resistant to chemical attack, and able to withstand high temperature.

Like so many other developments in the electronics industry, plant construction is almost entirely a matter of mechanical engineering, as Figs. 1 and 2 illustrate. Looking at the achievements since valve manufacture started after the First World War, I regard factory plant design as the hub from which have stemmed improvements in productivity, quality and quantity, the introduction of modern technological processes and the inspiration for achieving the almost impossible in the complex chemical and vacuum processes that are involved.

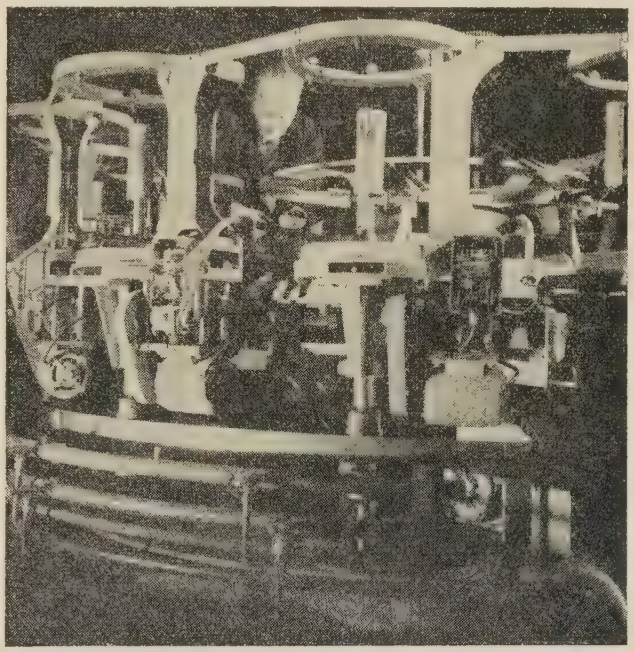


Fig. 1.—Constructing a 12-head cathode-ray-tube sealing machine.

The extent and nature of plant design and development problems are well illustrated by the fact that the tendency has been towards miniaturization in receiving and other valves, while the exact opposite is the situation with cathode-ray tubes. With the advent of the transistor, still smaller products have had to be manufactured, as shown in Fig. 3. The mere mechanical handling, during the various manufacturing processes, of these extremes in size is of itself a major engineering design problem, requiring in each case a quite different approach and a vastly different type of mechanism. Incidentally, added complications in cathode-ray-tube manufacture are the change from a circular- to a rectangular-faced tube and the introduction of metal backing of the luminescent screen.

However, these are not the only factors to consider in the design of production machinery, for, as already mentioned,

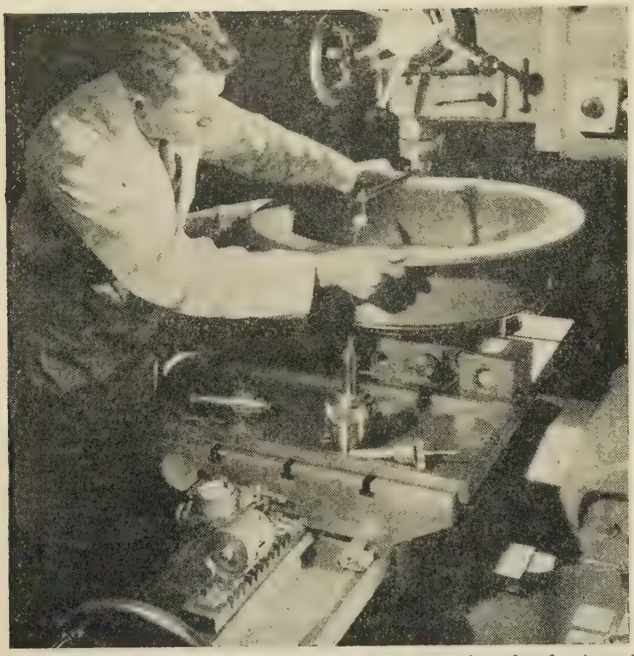


Fig. 2.—Setting up a 24-head turret in preparation for boring the indexing holes to an accuracy of 1/10 000in.

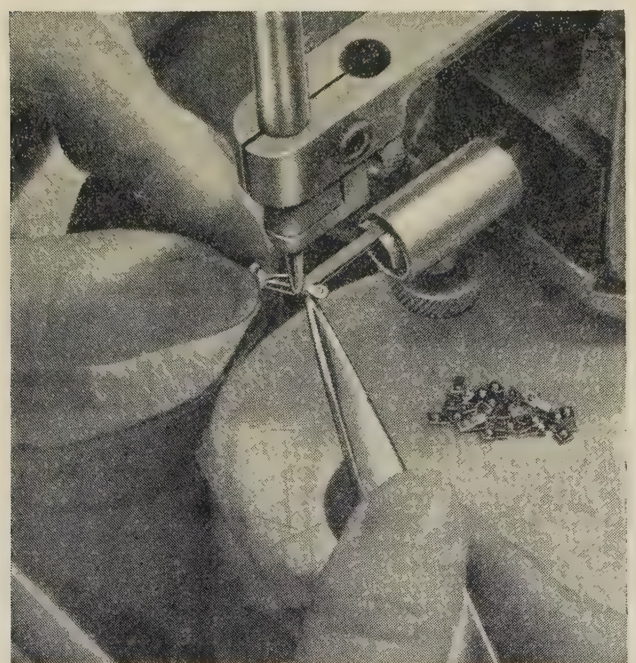


Fig. 3.—Welding a transistor assembly to its base.

manufacture involves the preparation of a vast number of different types of wires, electrodes, mica distance-pieces and glass tubing in various forms, all made to quite small tolerances, chemically cleaned and ready for assembly to form a complete valve which will eventually be exhausted, sealed and tested. Of major importance is the co-ordination of all the various pieces of plant in regard to speed and output, the avoidance of unbalanced stocks at any given stage in the manufacturing processes, and a smooth, unhindered, flow during the whole course of manufacture. Finally, it is in the nature of any vacuum device that, once made, it is impossible to alter its internal structure, which determines its characteristics, should it fall below the test specification or in any other way give an inferior performance in practice.

It will be of interest to consider the magnitude of this relatively new industry. In 1956 the turnover of the British valve industry was £25 million, and it produced 64 million valves of all types and over 2 million cathode-ray tubes, for radio and television, radar, communication, export, industrial applications, and to meet Government and Service needs. Additionally, a large number of the more specialized vacuum devices were supplied to various users for television cameras, X-ray equipment and instrumentation for nucleonic control. The corresponding numbers for 1949 were 19 million valves and 310 000 cathode-ray tubes, and this comparison gives a measure of the expansion of the industry in the post-war years. The relative expansion since 1930 is shown in Fig. 4.

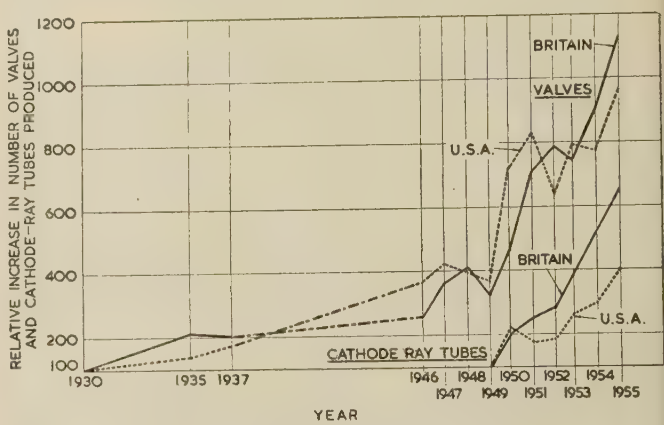


Fig. 4.—Growth of valve and cathode-ray-tube production in Great Britain and the United States.

It is difficult to obtain exact figures for the whole electronics industry, but those available suggest the impressive total of £300 million turnover for 1956, divided between professional equipment, entertainment equipment and vacuum devices as follows:

1. Government, industrial, communication, export (excluding valves):	%	%
Complete equipments	53	
Loose components	8½	61½
2. Entertainment equipment (excluding valves)		30
3. Valves—i.e. transmitting and receiving valves, cathode-ray tubes, thyratrons, photo-electric cells and other vacuum devices		8½
		100

It is significant that professional equipment accounted for a larger share of the total than entertainment items—a trend which will surely continue.

It is always interesting to compare the British and American valve industries. In the year 1956, measured in terms of receiving-valve output, the American valve industry was a little more than seven times as large as our own, produced 5½% of total United States industrial production, and accounted for 67% of the world production of 743 million valves; the British valve industry contributed 4½% of total United Kingdom industrial production and accounted for 9% of world valve production. Fig. 4 shows that the rates of increase of the American and British valve industries are remarkably similar.

There is no doubt that the general trend towards increased production will continue. Looking into the future we have to take into account the manufacture of transistors, which has now commenced here and in other countries, and which will continu

a rapidly increasing rate. I estimate that by 1960 the world oduction of valves and transistors will possibly reach a total 1 000 million.

THE CHANGING SCENE—NEW DEVELOPMENTS

Change has been the essence of valve development from the ne of Fleming's early work, and as our knowledge of electron nission, vacuum techniques, circuits, and so forth, has expanded, has the thermionic valve become increasingly a device of most limitless scope opening up an ever-widening horizon of plication possibilities.

It is obvious, therefore, that a great volume of research and velopment has always been undertaken in the valve industry, d as I cannot possibly mention more than a small fraction of e whole, I have limited my remarks to a few specific items. It ould be inappropriate in this Address to become involved in tailed technical considerations of these developments, for, like l aspects of electronic science and engineering, the technical ork involved is of quite a specialist nature, and they are sentially subjects standing in their own right.

Co-ordination of Valve Development

I will first refer to microwave valves, special-quality valves and mi-conductors; as evidence of the importance of these, during 47–56 over 100 articles were published on valve reliability d hundreds on transistors, and in the spring of 1958 The stitution is to hold an International Convention on Microwave alves, lasting several days. The reason I have chosen these ree is the interest displayed in them by the Government, for tional needs, through the agency of the Inter-Services Com-ittee for the Co-ordination of Valve Development (C.V.D.). ery wisely, arrangements were made during the war, and have ntinued ever since, to set up under the auspices of the Royal aval Scientific Service of the Admiralty a department which s become known as C.V.D., to foster and take the responsi-ility for valve developments for applications in equipment used the Armed Forces and to make sure that the various Services ould not be competing for limited technical resources in the lve industry. Since that time the C.V.D. organization and ort has grown rapidly, and even since the end of the war the ort applied by industry to Service needs has increased very any times. The present C.V.D. expenditure on development ntracts must amount to a few million pounds per annum.

Besides placing development contracts, C.V.D. organizes fective liaison with the valve industry and maintains close uch with development results, and has built up a balanced ttern of development projects, the sum total of which has ded considerably to our general knowledge of the three fields which I have referred.

C.V.D. sponsorship of these projects has undoubtedly been a ajor contributing factor in keeping this country in the van of ogress, and the foresight of the Government in setting up V.D. has paid handsome dividends, fully justifying the steps at were taken. The present occasion is an appropriate one which to record the valve industry's appreciation of the work C.V.D. and its personnel, and I would particularly like to take is opportunity of recording the contribution made by Sir ederick Brundrett. He it was who, in the early 1930's while Signal School, an Admiralty radio research establishment, w the necessity of inter-Service co-ordination of valve develop-ents, linked with the valve industry. It was from this con-ption that C.V.D. evolved and became one of the finest amples of inter-Service and industrial co-operation during the r and since.

Microwave Valves

Turning now to microwave valve development, a requirement of some importance is the availability of cathodes having extremely high emission current-densities, and a good deal of work has been directed towards the solution of this problem. As a result, some novel forms of cathode have appeared, one example being the dispenser type pioneered by Lemmens, where the emissive materials percolate to the surface of a porous tungsten body when the cathode is heated. This development has materially assisted the progress of microwave valve design towards meeting the ever-increasing demands for higher frequencies and powers.

Magnetron developments derived great impetus from military requirements for high pulse powers in radar systems. Initially conceived in this country by Randall, Boot and Sayers at Birmingham University, this device was rapidly taken up by America and became the subject of intense study. The main application for magnetrons is still in pulse radar and to a large extent will remain so, although some types of radar are likely to forsake the magnetron for devices such as the klystron or the travelling-wave tube.

In recent years the magnetron has made possible many new advances in conventional radar systems. The success of the British marine radar manufacturers in supplying the major part of the world's demand is well known, and becoming more common is civil-aircraft airborne radar for cloud and collision warning sets. The availability of magnetrons operating on wavelengths of a few millimetres has made feasible the radar control of ground movements on airport runways. Although radar using the pulse magnetron is considered to be rather old-fashioned, there is still a great deal of scope for expansion; in particular, one would expect to see magnetrons developed for still shorter wavelengths to provide greater resolving power for radar sets in cases where range is not important.

Apart from radar applications the magnetron has been extensively used as a source of radio-frequency power for linear accelerators. These instruments, in which electrons are acceler-ated to high energies, were initially developed as an atomic energy research tool and were then used for medical, radio-therapy and X-ray purposes. They now fulfil an important application in industrial processes connected with plastics. An example of this is the irradiation of polyethylene to increase its heat-resisting properties. The electrical industry is indeed indebted to the plastics industry for some of the modern advances in insulating materials, and the use of linear accelerators is, as it were, something given in return.

A limited application for magnetrons is as a source of power for radio-frequency cookers, and we may well see this gaining favour in the future. Food either pre-cooked or raw is placed in the 'oven' and subjected to strong microwave radiation, so being heated or cooked extremely rapidly by dielectric loss. The cost involved, however, may for the moment exclude its domestic use, but large caterers may favour the idea and use such a system.

The klystron was developed in its reflex form as a local oscilla-tor for superheterodyne receivers, and although the original ideas were American, a great deal of work was initiated in this country during the war, the contribution of Sutton and his team at Bristol and the effort at the Clarendon Laboratory being well known. Of late many improvements have been made in reflex klystrons in regard to ruggedness, reliability, life, and use at higher fre-quencies. Types have already evolved operating at wavelengths of 2 or 3 mm.

In the past the amplifier klystron received little attention because the magnetron produced radio-frequency powers with greater electronic efficiency and in return for far less development effort. Now, however, the situation is changing with the intro-

duction of phase-coherent radar, where the reflected pulse from a target is compared in phase with that which was transmitted, the resulting beat frequency being a measure of the target velocity. A fixed object will produce a zero beat in a phase-coherent system, and this identifying feature enables permanent echoes to be eliminated from the display screen. In addition, one can measure with great accuracy and instantaneously the velocity of a target—a point of considerable interest in any navigational problem, military or civilian.

There are many different variants of the phase-coherent radar system, an important common factor being the need for improved microwave amplifying valves, with the result that in this field work is being directed towards larger powers, higher efficiencies and increased bandwidths. Pulse powers of the order of megawatts are now possible and, in the lower-power types, gains in excess of 100 dB have been achieved. But by no means are all the answers known. Active research continues in the field of multi-cavity klystrons, where the problems are largely fundamental, involving the formation and maintenance of electron beams of very high current density. In slightly different applications the amplifier klystron has recently come into extensive use in navigational aids and in television broadcasting in Bands IV and V.

The travelling-wave tube was a most significant war-time invention, first conceived by Kompfner while working for the Admiralty. Kompfner's work showed that when a wave was travelling slowly along a structure such as a helix, a beam of electrons travelling at almost the same velocity and in close proximity to the wave would interact with it and amplify it, the extra energy in the amplified wave being derived from the kinetic energy of the electrons which are slowed down in the process. An important feature of the travelling-wave tube is the fact that, by special treatment of the electron beam after it leaves the cathode, the noise due to random velocity and density of the electrons can be largely removed, thus giving tubes with low noise factors which are particularly suitable for use as signal amplifiers.

The travelling-wave tube has received concentrated attention during the post-war years because the high gains and broad bandwidths obtainable make it ideally suited for use in centimetric wavebands for radio relays. The use of these wavebands has been necessitated by the big increase in the volume of traffic in telegraphy, telephony and television; for instance, the last named utilizes a bandwidth which could accommodate a thousand telephone conversations. The world's first radio relay using a travelling-wave tube was set up by the Post Office early in 1952 between Manchester and Kirk o'Shotts, since when complete ranges of tubes are becoming available for use in the programme for radio relays mentioned by Sir Gordon Radley in his Presidential Address of 1956.

Travelling-wave tubes are not necessarily limited to the few watts of power required in communication. Studies at really high power levels are progressing, and it may be that these tubes will become serious competitors of the amplifier klystron, for they are less complicated in many respects, having no tuned cavities and a simple magnetic focusing system. I believe that, in the long term, some communication will be carried out by the propagation of millimetric wavelengths along small-diameter waveguides. The possibility has already been demonstrated over fairly long distances, but many details remain to be worked out in connection with the actual propagation. In the meantime the valve industry realizes that the successful development of such a scheme is dependent upon the provision of practical oscillator and amplifier tubes, on which work is already proceeding, together with the more difficult and fascinating problem of miniaturization.

Backward-wave oscillators and amplifiers are particular forms of travelling-wave tubes having the advantage that they can be tuned by merely changing the anode voltage. It is difficult to see what the application of backward-wave oscillators will be, but apart from certain military possibilities one can imagine them eventually replacing the klystron as a local oscillator.

The International Convention on Microwave Valves will be of considerable importance in placing on record the huge volume of work that has been achieved in this field, not only here but in other countries. The gap in the spectrum between electromagnetic waves and infra-red radiation is closing, and how far this is so may be revealed in some of the papers to be presented.

Special-Quality Valves

During the period between the two world wars the urge for improvements in valve quality, performance and reliability resulted largely from the demands of the broadcast receiver manufacturers, and fruitful collaboration between them and the valve designers determined to quite a considerable extent the direction and scope of valve development. As a result, with the arrival of different applications, such as communication and airborne radar, the valves incorporated in them initially were those based on the design requirements of the domestic market. Such valves, however, were not primarily intended to withstand the shock and vibration associated with these new applications, so that there arose a need for a series of valves whose characteristics would not be impaired, nor their life shortened, by these more arduous conditions. The valve manufacturers, with the support of C.V.D., set to work to design valves to meet these new requirements, and found it necessary to study the properties of materials, the fatigue behaviour of glass, metal and mica, the influence of static stresses within the valve, the effect of cyclic stresses applied externally, the causes of gradual deterioration of electrical properties of the valve, and finally to review critically existing manufacturing methods. All this, and the development of sub-miniature types for missile work, has added greatly to our general knowledge of valve-manufacturing technology, the result being that several types of special-quality valves are now available, produced in air-conditioned plants under conditions similar to those obtaining in the manufacture of pharmaceutical products.

Semi-Conductors

The transistor was invented in 1948 at the Bell Telephone Laboratories in America during a study of the effect of external fields on the conductivity of semi-conductors. The first transistor was a point-contact type, but theoretical considerations soon showed that the principles could lead to a sturdier device—the junction transistor. The first prototype of this kind was made in 1951.

Because the transistor is an amplifying device, it has a very wide range of applications in the electronics field. It is extremely small and requires very little power to operate it; for these reasons it is destined to have a profound influence on the design, for example, of computers, of Post Office equipment such as repeaters and exchanges, and of Service equipments of every kind for use on land, sea and in the air. For the reasons already mentioned, the transistor will also greatly modify the design of components.

One of the main lines of transistor development, both here and elsewhere, is towards types which can operate at the higher frequencies. In America types operating up to 10 Mc/s and over are already in production; the theoretical limit, speaking from our present knowledge, seems to be in the region of 1 000 Mc/s or at the most a few thousand megacycles per second.

Another semi-conductor device with an important future is

ased on the fact that, when charge carriers are injected into a nction, multiplication occurs, rather like electron multiplicaon in an ionized gas. This effect can be used to produce veryigh-frequency oscillations, and from the theoretical point of ew perhaps eventually even in the region of centimetric waves.

The change between two possible energy states in molecules in be used for microwave oscillation or amplification, the intertion between a microwave system and a substance falling from higher to a lower energy state enabling power to be extracted. ne class of amplifier of this kind uses semi-conductors and ovides *m*icrowave *a*mplification by *s*timulated *e*lectromagnetic diation, from which is derived its abbreviated name of Maser. nis device can operate only at low power levels, and theoretically s efficiency and bandwidth are not great. However, useful iin can be obtained, and at the extremely low temperature at hich it must be operated its signal/noise ratio is extremely high.

The ideas using solid materials for electronic devices, which have outlined, are still so new that their importance and the tent to which they will be developed cannot be foreseen at esent. It is possible, however, that they may provide the key the whole field of millimetric waves.

Not the least important aspect of the great developments in is field is the enormous stimulus that has been given to research to the physics of solids. This field is growing in both volume id importance, and because of this we can undoubtedly expect is new science and technology to produce many novel and obably revolutionary devices and applications within the next w years.

Automation

Automation is no new thing and has been introduced in varyg degrees into many industries ever since the industrial revoluon. The reason, I believe, why so much attention has been cused on it in recent years is that, whereas formerly we made achines which could assist our physical efforts, we are now, rgely through the application of electronics, able to make achines which can augment our mental efforts. We can give achines 'electronic eyes and ears', to use the Duke of Edinrgh's expressive phrase, and these sensing elements can feed ck the information they receive so that any departure from the quired result can be corrected. In addition, it is now possible give a machine a programme which instructs it to perform rtain operations or to react in certain ways. In complicated stems these relatively new elements—feedback and programng—lead to the use of computers and electronic memories.

It is, in my view, the social rather than the technical implicans of automation which have attracted most attention. The o are, of course, interlinked, for if social conditions hinder the option of automation, then its technical progress and applican are bound to be retarded. This aspect is a very broad and eresting subject which I will not refer to any further here, ce I shall be opening an Informal Discussion on it before The stitution later this month.

Probably no other recent advance has had so many words ured out on it and has been so misunderstood. From what e reads and hears one could well imagine that it is only a atter of time before all our thinking and manual work will be rried out by machines. This is very far from the truth, and I l it would be appropriate for me to attempt to put this new velopment into some perspective.

Within a volume the size of our heads we each have a brain ighing some 3 lb and consisting, I am told, of some 10 000 llion individual cells,* which is about thirteen times the world oduction of valves in 1956. Among its countless other ctions, our brain includes the equivalent of a compatible black-and-white and colour television system, a sound recording and reproducing system, and an ability to recognize complex patterns which outstrips any practical mechanical or electronic equipment. If it were possible to construct a machine able to perform the same functions as the human brain, it would inevitably have to be largely electronic; if we brought together all the necessary component parts and could then in some miraculous way solve the vast problem of connecting them together, we should still be faced with the fact that, even with the most reliable modern components, several hundred would be faulty at any given instant. On whether a machine is capable of having original ideas, I would comment that, without vision there can be no progress, and however far I look into the future I cannot conceive that any machine man may create will ever be able to replace the relatively few geniuses, such as Leonardo da Vinci, Shakespeare, Beethoven, Wren, Kelvin and Faraday, on whose vision and creative ability the evolution of our civilization depends. On the day man is content to leave all his imaginative thinking to machines he will be destined to a future without beauty, without hope, and finally without love.

* BOWDEN, B. V.: 'Faster than Thought.'

Entertainment: Radio and Television

In general, entertainment radio comes in for a great deal of criticism, and we should remember that it is on the enterprise and success of this section of the radio industry that the valve industry was founded. But for the existence of a large entertainment radio industry, with its know-how and manufacturing capacity, radar sets and other Service equipment could not have been produced in the necessary quantities at the start of the last war, and who can say what might then have been the outcome? Without a large entertainment radio industry at home we could not to-day market competitively oversea.

Turning now to television, if we consider either black-and-white or colour, we find in both cases that the major development efforts centre on the viewing tube. For black-and-white we now have the introduction of tubes with electrostatic instead of electromagnetic focusing, with the consequent advantages of reduction in weight, easier setting-up and little or no variation in focus with changes in mains voltage. The other principal alteration is the increase in cone angle, which in turn requires wider-angle deflection. Just after the war this was 55°, subsequently it was 70° and production is now in transition from 70° to 90°. A further increase from 90° to 110° is the next logical step, and can be achieved with little increase in scanning power by bringing the deflection coils nearer the electron beam through a reduction in neck diameter; this modification will necessitate a complete redesign of the electron gun. On a 21 in tube, increasing the angle from 70° to 90° shortens the tube by 3 in, and a further increase to 110° reduces the length by another 3 in. This permits the use of a smaller cabinet, which is more acceptable in most living rooms and is also cheaper. Because of these advantages there is no doubt that the trend to wider angles will continue.

While there have been a few non-entertainment applications of colour television, notably in the field of medicine, for general entertainment the main problem remains that of finding a means of viewing far cheaper than anything which has yet been put forward. Cheaper tubes have been suggested, but they invariably require more complicated and thus more expensive circuit design. The development of a suitable tube is proving a remarkably intractable problem which is all the more conspicuous because in these times we have become accustomed to the rapid solution of nearly every technical difficulty. Some solution will, of course, be found, since it is not in our nature to relinquish such problems unresolved, but it may well take a few years yet.

Possibly the final solution will differ greatly from what we might expect with our present knowledge, for I have often felt that the cathode-ray picture tube, not only for colour but also for black-and-white, is a rather cumbersome and inelegant device, and in 10 or 20 years from now we may well look back on it as we now look back on the spark transmitter and coherer.

Low-Temperature Devices

In its onward advance, electronics is constantly establishing links with fields which only a few years ago seemed quite remote. In recent years it has reaped much by its association with ferrites and semi-conductors, and now it is once again crossing the boundaries of new territory—this time into the low-temperature field. There are three main reasons why this advance is taking place; firstly, recent progress in low-temperature techniques has made the creation of temperature near absolute zero a matter of no very great difficulty; secondly, more complicated and therefore larger equipments costing, say, £100 000 or more, in which low-temperature facilities can economically be provided, are nowadays more frequently considered; thirdly, because of the extremely low noise level at temperatures near absolute zero, devices which operate in this region hold great attractions for the electronic engineer.

There are two main developments in low temperatures—I have already mentioned the first, the Maser, in dealing with semi-conductors. The second, the Cryotron, is a switching device which consists of a straight wire with a second wire coiled round it, both immersed in liquid helium and therefore having effectively zero electrical resistance. However, the resistance of the straight wire is restored when a sufficiently strong magnetic field is produced by a current in the coiled wire. The current required is only a fraction of that in the straight wire, and the effect can be produced within microseconds; the device thus acts as a very-high-speed relay. It has the advantage of simplicity and small size, and although it is too early yet to judge its value fully, it may well have wide applications in large computers.

Frequency Allocation

Frequency allocation is a problem somewhat different from those to which I have so far referred. I mention it here because of the supreme importance I attach to the use of what is an irreplaceable raw material, the ether, which in the fundamental nature of things is a strictly 'one-off' item. The growth of radio-communication, broadcasting and navigational aids has made the problems of frequency allocation more and more acute, especially during the last decade. International agreement has on the whole been good considering the complexity of the subject; and this is as well since the problem of interference has recently acquired a new facet with the advent of scatter transmissions, because these can cause interference here even though neither terminal is situated in this country. Somewhat in contrast to the international scene, some of us have doubted whether allocation at national level has taken the needs of all users sufficiently into account. Allocation to suit everyone is, of course, an extremely difficult task, requiring as it does the co-ordination and reconciliation of the needs of many diverse users and often involving security considerations. For all these reasons we should welcome the Postmaster General's statement in the House on the 3rd July, 1957, that he intends to set up a Frequency Advisory Committee to guide him on the broad aspects of frequency planning. I do not think I am revealing any secrets if I say that these proposals are very much in line with ideas which the Radio Industry Council discussed on a number of occasions with the previous Postmaster General.

EDUCATION AND TRAINING

This is the age of unprecedented advance on all fronts of applied science, when so much depends upon the skill and ability of the individual and when the nations of the world, almost without exception, are acutely conscious of the need to harness their man-power and material resources effectively in the struggle to maintain and improve their standards of living, to say nothing of their desire for international prestige and national security.

It is not surprising, therefore, that a good deal has been written in the past few years on education and training, pointing out the urgent need for a great increase in the number of adequately trained scientists and engineers. Some progress has been made, but the problem is being continuously reviewed from many aspects, for it is realized that our future progress industrially and therefore nationally depends upon the work of individuals of ability possessing the requisite learning and skills to enable them to sustain the immense activity necessary for the creation of new knowledge and the development of new technologies.

The importance of the individual in this respect needs emphasis, particularly in these days of increased mechanization, such as the development of automatic control techniques, data processing and many other labour-saving devices, for there is a danger of giving too much prominence to these things in the mistaken belief that man-power in consequence becomes less important. The contrary is the true position, however, for history shows that, in the long term, increased mechanization has not by any means led to increased unemployment. I do not intend to pursue this subject further, but I want to use this opportunity to suggest some lines of action which I believe to be necessary if we are to be successful in training a much greater number of suitably qualified technical personnel in a reasonable time and to the standard demanded by present-day conditions.

These conditions demand that the training of the potential scientist should include sufficient liberal studies to make him aware of and appreciate the social and economic problems of the modern world. At the same time, the potential classics man must acquire more scientific knowledge of a general character and so be able to appreciate the effects of scientific advances on modern conditions and to know something about the elements of science.

The question naturally arises: In what way can this be achieved? One suggestion is to raise the level of science teaching for all those in secondary schools, which, as I think is generally conceded at the present time, is on too low a level. Further, this teaching should commence at an earlier age than is the case at the moment, so enabling students to proceed to more advanced work at an earlier stage in their school life, at the same time assisting them to reveal and develop natural talents and preferences. However, we must not forget that in secondary schools the success of science teaching, even at the present level, is entirely dependent on the provision of properly trained teachers. The inadequate supply of suitable candidates for the teaching profession is a fundamental problem, and unless we can solve it the penalty will be severe and enduring.

Inevitably these suggestions raise the question of the amount of time available, and a solution of this problem may mean an extension of the period spent at school. The achievement of this objective may also require that more attention be given to 'non-faculty' subjects in assessing university entrance standards, which in turn would involve changes in the present accepted G.C.E. subjects at the advanced level. Some practical support to this line of thought would be given by the suggested Cambridge Honours Course combining science and the arts, which would also give encouragement to the sixth-form boy on the arts side to study science.

Referring to teachers for technical colleges, the report, 'The

pply and Training of Teachers for Technical Colleges', of a ecial committee under the chairmanship of Dr. Willis Jackson pointed by the Minister of Education in September, 1956, and blished a few months ago, sets out very completely the nature d extent of the problem. Of special significance to this stitution and to the whole electrical industry is the statement future needs, showing that a full-time teaching staff of 11 500 1955–56 should be increased to 18 600 in 1960–61, and part-ne staff from 39 000 to 47 000 in the same period. Equally gnificant is the opinion expressed by the committee that ucational institutions and industry tend 'to regard themselves separate and distinct rather than as intimately associated rtners in a joint enterprise'.

In spite of the many difficulties in its application, industry ould most carefully consider how to implement the following atement appearing in this report: 'Industry . . . must be willing accept and indeed to encourage and assist the transfer to ll-time teaching work of experienced staff it can ill afford to se as the only means of ensuring a much augmented supply of itable recruits of high quality'.

Before leaving this topic, I would suggest to the electrical dustry the need for a careful assessment and understanding of e implications of industrial training as part of the sandwich urses now being approved by the National Council for Techno-gical Awards. These courses will be run by approved technical lleges and will lead to the Diploma in Technology (Engineering). ne student gaining this award is assumed to be knowledgeable both the fundamentals of science and technology and also in e practical application of this knowledge. It follows, there-re, that the purely academic training, and the technological aining carried out in industry, must be treated as of equal nportance in order that the award may denote an adequate andard of academic achievement and of competence in techno-gical application. The danger is that the periods of industrial aining may tend to be regarded as periods of normal employ-ent unrelated to the academic part of the sandwich course. ould this become the general pattern, then the objective of oducing technologists, that is to say engineers with experience the practical application of fundamentals, will not be achieved. is my considered view that industry must take an active part determining the standards and requirements of the techno-gical training and must insist on these standards being aintained.

The new course for the Diploma in Technology is only one of any factors which are increasing the demands for practical aining, the burden of which is carried almost exclusively by e larger industrial concerns. There is a limit to what these ganizations can do, however, and even though this limit has erhaps not yet been reached, it is clearly right and necessary at smaller firms should contribute their share to the pool of actical training facilities. A very few such firms are already ntributing through co-operative training schemes, and they ave shown that, given the right conditions, such schemes can e successful. Because extra men will be available owing to the ulge going through the secondary schools, and because of the il-off in call-up, there is an urgent short-term need to provide ditional practical training facilities, quite apart from the long-rm implications of the position. It would be a disastrous aradox if, at the time when schools, colleges and universities e increasing their outputs in an attempt to provide the addi-onal men which industry has called for over such a long period, dustry itself should fail to have available the practical facilities ithout which training cannot be properly completed.

We must all welcome the Government decision for a progressive crease in the numbers called up for National Service, which is w likely to end in 1960; in addition, it is probable that in 1958 and 1959 not all those liable for Service will actually be called up. However, in these two years all trained men who have been deferred will continue to be called up even if more are available than the Forces require to meet their needs for skilled man-power. Obviously, industry could well do with this surplus. This waste of our most valuable commodity—trained personnel—is very disturbing, and one must hope that the policy behind it will be further considered.

Among the many aspects of education industry has to deal with is that of keeping its own technical personnel of all grades up to date with the many new ideas which are constantly coming forward. Usually this can be done by reading and by outside lectures, but occasionally, when quite a new concept is intro-duced, something more than this is needed. Transistors, for instance, require the engineer, who is accustomed to working with valves, to think in quite a different way. Instead of thinking in terms of voltage he must think in terms of current, and practically all his accustomed habits of design thought must be discarded. In such instances industry has the task of re-educating its own engineers, and some special effort is required. In the case of transistors the need has been met, in one instance, by the production of a film which explains the theory of transistor operation and then proceeds to the derivation of the equivalent circuit; such an approach enables the engineer to base his new ideas on the solid foundation of basic fundamentals. As electronics expands ever more rapidly into new fields, the valve industry will have to give increasing thought to this problem of re-education.

Recently the recruitment of larger numbers of women to the ranks of technicians and professional engineers has been increasingly advocated. In a country such as ours, which still holds rather definite views on the scope of women's activities, such a course presents its own peculiar difficulties and requires a change in outlook among parents and teachers and in industry. It is time we ceased to fear the shadow which the wedding ring casts behind and before it. Already this fear is less than it used to be; for instance, marriage is no longer a bar to some posts, in teaching and in the Civil Service for example. But we in industry are changing our ideas far too slowly. In the long run, I believe, prejudices will be overcome; let us hope the process does not take too long, because we urgently need at this moment the skilled contribution which, I am sure, suitably trained women have it in their power to give. The reason we turn to them is quite simple; they are the most significant source from which we can meet our persistent and dangerous shortages of trained personnel.

THE PRESS

I am very pleased to include in this Address some mention of the Press, because The Institution and the electronics industry owe so much to the support they receive from both technical and lay papers.

It is my view that the public require to be still better informed on scientific achievements and that there is a great deal more to be done in educating them to a deeper appreciation of the significance of our work as engineers. The chief means of closing this gap between science and those who work in other fields is through the lay Press, and we are grateful to them for what they have already done in this direction and for bringing our achievements to public notice. I am convinced that the role they are playing in this way will be of ever-growing importance.

One of the difficulties with which engineers are faced to-day is that the rate of advance is so rapid that many of us tend more and more to become specialists in our own particular subjects, and develop an increasing ignorance of other work which has no

immediate bearing on what we are doing. It is the technical Press which plays a major role in bridging this gap; it is also to the technical Press that we look for the broad and intelligent exchange of ideas and information, and for the publication of the results of our individual work. I like to think of the technical Press as complementary to the journals of the learned and professional institutions, the whole constituting a shop front in which we can display our scientific wares, not only amongst ourselves but also to other scientists all over the world.

It is appropriate to acknowledge the accurate and painstaking work of the editorial staff of our own Institution. When you consider the wealth of papers, articles and other publications produced by our members, and when you remember that all this, and a good deal more, has to be sifted, assessed, edited and presented accurately and intelligently in the various Institution publications, then I think you must all agree with me that our editorial staff does a really magnificent job.

BEYOND THE HORIZON

I conclude my Address by leaving with you a thought which I personally feel is of growing importance, not only for us as members of this great Institution, but for all scientists and engineers in all the countries of the world, and which may well become the most significant issue that members of professional bodies such as our own have so far had to face. I refer to our role as leaders, not in the sense of fitness for an executive position, but as leaders who accept the responsibility for guiding opinion on the significance and implications of scientific engineering advances, so that these advances assist the lot of mankind rather than the reverse.

We have invariably regarded it as right that we should be concerned only with establishing scientific truths, putting new knowledge to some practical use, producing some commodity at the right price, at the right time and in the right place, and showing how engineering science can be successfully harnessed to fulfil our many needs. With a few exceptions we have been little concerned with the social and political repercussions of our efforts; the use which others make of our work has far too seldom appeared to us to be of prime importance. I believe that the time will soon be here, if it has not already arrived, when more of us must be as much concerned with the results of our scientific and engineering achievements as we are with the work itself.

For instance, we have made automation possible—what is our responsibility for the social changes it will bring? Lately there has been considerable public discussion on the value of television; it is said that on these programmes some £18 million per year are spent by the B.B.C. and I.T.A. together. One cannot help reflecting whether the engineers who created this medium have not some special responsibility for the way in which it is used. They cannot stand aside and make no contribution whatsoever to current thought on these matters. The scientist must take part in government at the highest level, as exemplified by Sir Winston Churchill's appointment of the late Lord Cherwell as his scientific adviser during the last war.

However, it is in the military sphere that this question of our responsibility has appeared in its most acute form, and it is not a new problem by any means. H. D. Smyth of Princeton University, New Jersey, has admirably stated the situation in the first report ever published on the military use of atomic energy. Writing in 1945, he said, 'We find ourselves with an explosive which is far from completely perfected. Yet the future possibilities of such explosives are appalling and their effects . . . on international affairs are of fundamental importance. Here is a new tool for mankind. . . . Its development raises many questions that must be answered in the future. These questions are not technical questions—they are political and social questions and the answers given may affect all mankind for generations.' We are very little nearer finding an answer to these questions than we were twelve years ago when this was written.

In the meantime we now have a completely new situation, and must face the fact that the world has been brought to a crossroads; one road leads to the destruction of world civilization as it now exists, the other to a better life for everyone, not only in backward countries but in highly developed countries too. It is we, the engineers and scientists, who are largely responsible for this position, and surely we are morally responsible for the way in which our work is used. Can we, or dare we, stand aloof from the social and political implications of what we have achieved? I am convinced that this question must sooner or later be given an answer. I am equally convinced that this great Institution of ours has a vital part to play in framing that answer.

The Institution of Electrical Engineers
Abstract No. 2508
Jan. 1958

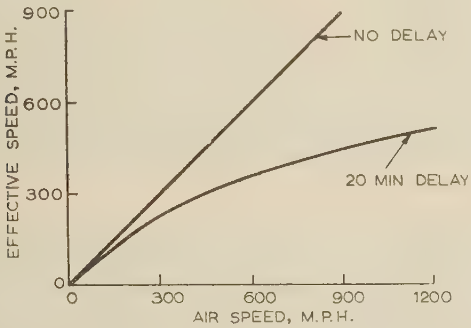
RADIO AND TELECOMMUNICATION SECTION: CHAIRMAN'S ADDRESS

By J. S. McPETRIE, Ph.D., D.Sc., Member.

'SOME RADIO AIDS FOR HIGH-SPEED AIRCRAFT'

(ABSTRACT *of Address delivered* 16*th October*, 1957.)

If an aircraft is delayed either in take-off or landing, the delay rings about a reduction in effective speed. Some idea of the agnitude of this effect can be seen from Fig. 1, for which



g. 1.—Loss of speed due to a delay of 20 min in an air journey of 300 miles.

as been assumed a 10 min hold-up at each end of a 300-mile ight. It will be noticed that the effect of this delay becomes rogressively greater as the aircraft speed is increased. As e cruising speed increases, therefore, it becomes ever more nportant to reduce to a minimum any delays en route.

Effective control of an aircraft near a departure or arrival rport requires adequate communication between the controller n the ground and the pilot in the aircraft. At present, the ain communication link to the aircraft for short-range working by v.h.f. equipment, which has now been in operation for ther civil or military aircraft for nearly 20 years. It is worth calling that this country in the last war was the first in the world evolve a system for the close control of fighters by the use of dio, and that this was an essential link in the defence chain erating during the Battle of Britain.

V.H.F. equipment is being replaced in all N.A.T.O. military rcraft by u.h.f. equipment operating in the 225–400 Mc/s freuency band. Although u.h.f. equipment, for the present, is nfined to military aircraft the facilities offered are of the type hich might be expected in all aircraft in the near future. In e airborne equipment the pilot can select no less than 1 750 parate channels simply by dialling the required frequency; in dition, he will have 18 preset channels and the distress channel, available on a pushbutton basis.

Because of the quasi-optical characteristics of radiation in the .f. and u.h.f. bands the ranges of such equipments are limited the horizon ranges of the aircraft. For longer ranges we must e equipment operating at lower frequencies in the h.f. band, 1–30 Mc/s. The system of ground-air communication which probably be designed for use in this part of the frequency ctrum will be of the single-sideband type described in a per by Mr. Barnes.* The incorporation of single-sideband techniques in h.f. equipment should give substantial improvement in long-range communication with aircraft over the existing systems which are of double-sideband type. The performance of airborne single-sideband communication may well be sufficient for all long-range ground-air communication to be made by radiotelephony. If this were satisfactory, a specialist signaller in the aircraft could be dispensed with, thus resulting in an increase in payload; at the same time, by communicating directly by speech with the pilot, both the handling time of any message and the probability of error in interpretation would be greatly reduced.

However adequate the communication between the ground controller and the pilot, the instructions to the latter cannot be satisfactorily carried out unless he already has a fairly precise knowledge of his position relative to the track which he is expected to follow. Any deviation from this track, in addition possibly to endangering the aircraft, will cause additional journey time, which is equivalent to a reduction in effective aircraft speed just as much as a delay at either end of the journey. The pilot must therefore have adequate navigation information at all stages of the flight.

There are a number of air navigation aids which have been in use for some time, such as D/F, Radio Compass, V.O.R., D.M.E., Gee, Loran, Decca and Console. As most of these have been described before, remarks will be restricted to some of the newer radio-navigation aids which might be expected to come into general use in the next few years. Typical of these newer aids are Tacan, Doppler, Dectra and stable-frequency oscillators.

As its name implies, Tacan was developed originally for tactical use by single-seat fighters. The interpretation of the results obtained, therefore, had to be as easy as possible, any complexity being relegated to the design of the circuits in either ground or air equipment. Essentially, Tacan consists of a ground beacon operating in the 1 000 Mc/s frequency band and energizing a cylindrical aerial system rotating at a speed of 900 r.p.m. This aerial gives rise to a modulation in the signal received by any aircraft tuned to the beacon, and the phase of this modulation as measured in the aircraft is sufficient to determine the direction of the beacon from the aircraft. The ground beacon incorporates a transponder which can be triggered by the aircraft transmitter, the distance between the aircraft and beacon being given by the time between the initiating pulse from the aircraft transmitter and the reception of the signal in the aircraft from the beacon transponder. The duty cycle of the ground transmitter remains constant by ingeniously modulating it with noise pulses while no aircraft is interrogating the beacon and replacing the noise pulses by the interrogation pulses until the beacon becomes saturated with 100 simultaneous aircraft interrogations.

The bearing and distance of the interrogated beacon are both displayed in the pilot's cockpit on a single dial similar to that of a car speedometer, the directional information being given by a centrally mounted pointer and the distance by an inset counter. The range, as would be expected for a system operating in the 1 000 Mc/s frequency band, is limited to that of the optical horizon of the aircraft, with an overall limit of 200 miles. Aircraft

* BARNES, G. W.: 'A Single-Sideband Controlled-Carrier System for Aircraft munication', *Proceedings I.E.E.*, Paper No. 1535 R, August, 1953 (101, Part III, . 1).

Dr. McPetrie is at the Royal Aircraft Establishment.

speeds of up to 1 500 m.p.h. can be accommodated. The Tacan system is not in general operation yet, there being only one ground beacon in the United Kingdom installed at Farnborough. However, there is every hope that before long its use will be fairly widespread, as it is the first navigational aid which gives directly both distance and bearing of an aircraft from a known ground position.

The Decca navigator has been extensively used for ship navigation and is becoming popular as an air-navigation aid. Owing to errors arising from ionospheric propagation effects its range is restricted to some 300 miles. A modification of the system to operate at ranges of 1 000 miles or more and called Dectra has recently been tested in a Valiant aircraft across the Atlantic. The results suggest that in Dectra we may have an extremely valuable navigation aid for aircraft on transoceanic routes. Briefly, the system consists of a master and slave station operating at a frequency in the region of 70 kc/s and located about 100 miles apart normal to, and at each end of, the route along which a series of tracks is required. The proposal is that Dectra and Decca should form an integrated system, some of the airborne equipment being common. On leaving an airfield the pilot would have all the facilities for accurate navigation using Decca. This would allow him to navigate his aircraft into the Dectra field system so as to take up the track which he was scheduled to follow. Each Dectra track, of course, corresponds to a given phase difference between the radiations from the nearer master and slave Dectra stations. Near mid-route the pilot would transfer to the appropriate track formed by the two Dectra stations at the terminal end of his journey. When he approaches these stations, he can then obtain more accurate fixes from the Decca chain covering his destination airfield.

It is proposed that range determination for Dectra should be given by comparing the phases of the radiation received in the aircraft from the master stations at each end of the system. It is envisaged that continuous reception from both ends of the route may not always be possible, and so a stable-frequency oscillator will be carried in the aircraft as a memory aid during periods of poor reception. At present, the only Dectra stations are located in Newfoundland. In spite of this limitation, however, the Valiant aircraft has made a number of successful flights between Newfoundland and the United Kingdom, in one of which the first land sighted was on breaking cloud over the home station at Boscombe Down.

Changes in relative phase between the radiation from a stable-frequency transmitter on the ground and the phase of a correspondingly stable oscillator in an aircraft at nominally the same frequency can be used directly to determine changes in radial distance between the ground transmitter and the aircraft. When associated with any system giving the direction of the aircraft relative to a known ground position the pilot can use his stable oscillator to determine his actual position. Alternatively if two stable-frequency ground transmitters are available, the distance of the aircraft from each transmitter can be determined and thus the position of the aircraft. This system has not been used to date because of the extremely high order of frequency stability which is required. This can be seen from the following Table,

$\frac{\delta f}{f}$	Error		
	10 min	30 min	1 hour
	km	km	km
10^{-7}	18	54	108
10^{-8}	1·8	5·4	10·8
10^{-9}	0·18	0·54	1·08

which gives the distance error resulting from comparing two oscillators supposed to be on identical frequencies, f, but actually being on slightly different frequencies, f and $f + \delta f$.

It will be noted that the frequencies must lie within 10^{-9} of one another to reduce the errors to the order of 1 km over a period of one hour. Since the error in range determination arising from a given error in frequency setting is proportional to time, the accuracy of the stable-frequency oscillator as a navigational device becomes correspondingly greater as the journey time itself decreases. We may expect, therefore, to see increasing use of such devices as aircraft speeds increase.

Research in both the United States and the United Kingdom is proceeding into methods for using the spectral lines of molecules such as ammonia and caesium for the control of stable-frequency oscillators. The accuracy of frequency setting for such oscillators should be adequate for this purpose. Their frequency has the unique advantage of showing no frequency drift, a characteristic of, for example, quartz oscillators.

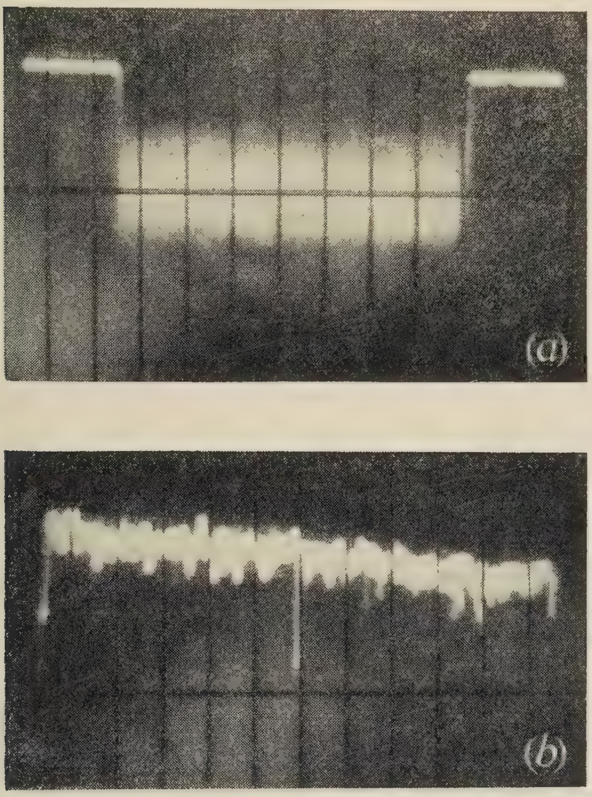
Most navigation systems require both ground and air equipments. Self-contained navigation aids are now making their appearance, and one of the latest of these is Doppler navigation. In this system, a radar carried in the aircraft illuminates the ground by forward and rearward looking beams. The ground returns in the two cases are compared in frequency, and as the difference corresponds to twice the Doppler component, the velocity of the aircraft relative to the ground can be determined. In the equipment produced in this country, the airborne radar also feeds alternately beams to port and starboard. By this means the effect of drift can be taken into account.

One great advantage of the Doppler navigator is that, theoretically, it has no limit in range. However, for the longer ranges it is advisable to check and correct, if necessary, the system when the aircraft passes over recognizable ground features en route.

The pilot is now well on his way towards his destination, having been given suitable instructions by radio, and navigating his aircraft with the assistance of, possibly, one of the navigation systems described above. When he reaches a distance of about 50 miles from the arrival airport, the pilot asks for a check of his position. This is obtained by determining his bearing relative to a number of ground D/F stations, the resulting fix being reported by radio to the pilot by the ground controller, who will also give suitable instructions as to how he wishes approach and landing to be carried out. The aircraft may be taken under control by the controller and talked down to eventual landing. Alternatively, the pilot may approach the required runway using his altimeter and the instrument landing system.

An aircraft is made to approach the destination airfield along a series of recognized routes. With increasing density of air traffic and, particularly, with increasing aircraft speeds, the air-traffic controller requires correspondingly increased range on his ground radar. However, increasing the range of a ground-surveillance radar also increases the amount of information in all directions as well as along the traffic lanes. It would be much better for the controller to have access to the information gathered by radars located at some distance along the traffic lanes as well as a more-modest-range radar close to the airport. Such an arrangement of remote radars brings in the problem of transferring the information gathered by them to the ground controller. There are various possibilities of achieving this objective; for example, the complete p.p.i. display could be transmitted by radar relay. However, this is much too inflexible and expensive and, at the same time, sends more information to the ground controller than he requires for air-traffic control purposes. Some work will be described which has been carried out at the R.A.E. by Mr. Hinckley on the automatic processing and transmission of radar data over normal telephone lines. In this

heme, the 'raw' data from a plan-position search radar under-oes signal integration, thereby allowing automatic detection of rget signals without appreciable loss in maximum range of etection. Some idea of the gain in discrimination of the signal om noise can be obtained from Fig. 2.



ig. 2.—Gain in discrimination of signal from noise by means of signal integration.

(*a*) Input signal immersed in noise approximately 1 : 1.
(*b*) Output signal showing improvement after integration of 10 inputs.

In the R.A.E. equipment, the occurrence of an integrated rget signal above a given threshold automatically causes the R, θ) co-ordinates of the target to be measured in binary digital orm. The appropriate (X, Y) co-ordinates are generated simul-neously and passed to a magnetic-core store. While the rate f input information to this store depends on the azimuthal istribution of aircraft, the information contained in the store is extracted at a steady rate such that the content of the store, if full, is extracted over a period equal to the rotational period of the radar aerial. The store, as at present constructed, can accommodate 22 digits of information about each of 625 aircraft. With a typical rotational period for the radar of 10 sec, the rate of transfer of information from such a store is 1 375 bits per second, a rate which can easily be accommodated by a normal telephone line.

All the relevant air-traffic control information at the distant radar can, by this means, be transferred automatically over a normal telephone line to the ground controller. As the information is already in Cartesian co-ordinates, allowance for the displacement of the remote radar can readily be made, so that, if necessary, the remote targets could be displayed directly on the plan-position indicator already in use by the ground controller, or in any other way.

The conventional p.p.i. display of a radar at the R.A.E. has been compared with the corresponding decoded display after integration, automatic detection, conversion from polar to Cartesian co-ordinates, digital encoding, storage, with subsequent transmission over 150 miles of normal Post Office telephone circuit which passed through six public exchanges, and final decoding and display. In spite of the complex radar processing involved, little information is lost; in fact, the processed display is better than the original because in the processing certain irrelevant information is automatically suppressed.

It is believed that some such method for the automatic processing and transmission of radar information will be essential in the near future at all major civil airports in order to accommodate the increase in traffic density and aircraft speed which is to be expected.

Incidentally, a subsidiary but very useful advantage of the radar processing scheme is that, as the information can be carried adequately by a normal telephone circuit, it can equally well be transferred to a magnetic tape or similar recording. A p.p.i. display, in fact, has been made at the R.A.E. from a normal magnetic tape carrying processed radar data. Such a recording could be made of each aircraft approach to, and landing at, an airfield, when information would be available for any subsequent accident inquiry.

Looking further into the future, we can envisage a time, not so far distant, when the processed radar data can be used in a ground-based digital computer and appropriate command signals sent automatically to the aircraft controls so that a fully automatic approach and landing would be possible.

In this Address only certain highlights in the use of radio and electronic aids to flying can be touched on. It is probable that the use of such aids will become ever more important as both the traffic density and speed of aircraft increase.

CENTRE AND SUB-CENTRE CHAIRMEN'S ADDRESSES

The Institution of Electrical Engineers
Abstract No. 2491
Jan. 1958

SCOTTISH CENTRE: CHAIRMAN'S ADDRESS

By E. OPENSHAW TAYLOR, B.Sc., F.R.S.E., Member.

'SCOTTISH ENGINEERS AND THE SCOTTISH ELECTRICAL TRAINING SCHEME'

(ABSTRACT *of Address delivered before the* SOUTH-EAST SCOTLAND SUB-CENTRE *at* EDINBURGH 1*st October, the* SOUTH-WEST SCOTLAND SUB-CENTRE *at* GLASGOW 2*nd October, and the* NORTH SCOTLAND SUB-CENTRE *at* ABERDEEN, 4*th October, and* DUNDEE, 17*th October*, 1957.)

This Address reviews some of the early contributions made by Scottish engineers to British industrial progress and discusses some aspects of present student education in Scotland.

About 250 years ago, at the time of the Union of Parliaments, Scotland was almost a feudal State with only a little shipbuilding, fishing and primitive coal-mining giving importance and employment to a few East Coast towns. About this time, however, the foundations of modern engineering were being laid in England by the development of the iron industry, and soon after, Dr. John Roebuck stepped on to the Scottish scene. Roebuck was trained in medicine at Edinburgh University, but subsequently, with the help of Samuel Garbett, a Birmingham business man, he entered the iron industry by building the Carron iron works, near Falkirk, in 1759. Whereas in England the iron industry had grown up gradually over a long period, in Scotland, due to Dr. Roebuck, it started suddenly and on a relatively large scale; the Carron works, which brought the word 'carronade' into the language, still prosper to this day.

Into this world was born the first of the great Scottish engineers, James Watt. In 1755, at the age of 19, he decided to become a mathematical instrument maker and, like many subsequent Scottish engineers, went to England for his apprenticeship. The journey took him twelve days over almost non-existent roads; returning to Glasgow as instrument maker to the University he interested himself in the steam engine. He devised the separate condenser, which, following Newcomen's idea of the separate boiler, so greatly increased the efficiency of the engine as to make it into the prime mover that introduced the industrial revolution into Britain.

It is important to realize that Watt did not suddenly come upon the idea of the separate condenser; he had wondered for a long time how to overcome the inefficiency of the Newcomen engine, and, setting an example to later students and engineers, he made a close study of the basic principles underlying the work he was attempting—he investigated the properties of steam. It was this fundamental knowledge that enabled him to visualize the solution to his problem when taking a Sunday afternoon walk across Glasgow Green. On Watt's death a monument was erected, by public subscription, in Westminster Abbey; in the Capital of Scotland a similar fund was inaugurated by Lord Cockburn which eventually resulted in a statue and in the Watt Institution and School of Arts; this Institution has developed into the Heriot-Watt College and the statue stands outside the present building.

William Symington of Leadhills also took out patents for a steam engine, leading to some discussion with Watt over possible infringements. One of Symington's engines was installed, in 1802, in the *Charlotte Dundas*, which sailed on the Forth and Clyde canal and was the world's first practical steamship. It is said that Robert Fulton, an American but the son of a Dumfries man, and Henry Bell of Linlithgow saw the trials of the *Charlotte Dundas* before developing the *Clermont* and the *Comet*, which achieved fame in their respective countries as pioneer steamships.

The deplorable state of the roads encountered by Watt on his journey to London was remedied largely by two men from Dumfries—John Loudon McAdam (1766–1836) and Thomas Telford (1757–1834). McAdam's major contribution was a new form of road surface comprising small stones laid to a thickness of about a foot on the natural soil, and so successful was his method that many miles were laid and he earned for himself the title 'Colossus of Roads'.

Telford in his young days worked as a stonemason in Edinburgh and subsequently built not only roads but also harbours, bridges, canals and tunnels and eventually became the first President of The Institution of Civil Engineers. Telford set an example to many later Scottish engineers by extending his activities outside Britain, a notable achievement being his construction of the Gota canal in Sweden. He also set another example that might well be followed by modern engineering students in that he was, on all occasions, well and correctly dressed.

Contemporary with Telford was John Rennie (1751–1821), an Edinburgh University student, subsequently employed by Watt. Among the numerous harbours and bridges that he built was the original Waterloo Bridge, which, although no longer in existence, is still regarded as one of the world's masterpieces.

Returning to mechanical engineers one finds the name of James Nasmyth, inventor of the steam hammer, who was born, educated and buried in Edinburgh. He started business as a millwright in Manchester in 1834 with one assistant, and, 22 years later, retired at the age of 48, leaving the large Bridgewater Foundry employing over 5 000 people and was able to enjoy 34 years of retirement.

One must not forget the contribution of Scottish engineers to the development of the steam locomotive, and among many names, two stand out as having had influence far beyond the borders of Scotland. Patrick Stirling was born at Kilmarnock in 1840 and rose to be Locomotive Superintendent of the Glasgow and South Western Railway before going to the Great Northern in 1866, where he built the famous single-wheelers that hauled the East Coast expresses of the day at speeds up to 80 m.p.h. Engine No. 1, still preserved at York, took part in the great Railway Race of 1895, when the journey from King's Cross to Aberdeen was done in 8½ hours. The other name is that of Dugald Drummond, who became Locomotive Superintendent of the North British Railway, of the Caledonian Railway, and finally of the London and South Western Railway. Drummond engines were characterized by their simplicity and robustness, but, nevertheless, one of his engines, on the last night of the Railway Race, hauled the West Coast express over the difficult

Mr. Taylor is at the Heriot-Watt College, Edinburgh.

aledonian section of the route from Carlisle to Perth at an erage speed of 60 m.p.h.

It was while these spectacular achievements in mechanical and vil engineering were taking shape that Michael Faraday was ying the foundations of electrical engineering. The possibility obtaining an appreciable and controllable force led several ientists of the time to devise simple electric motors. Robert avidson of Aberdeen seems to be the only Scottish participant the field at this time; by 1839 he had a lathe and a truck perated by electric motors; the Royal Scottish Society of Arts oted him a small sum for his experiments, which culminated in 16 ft battery-driven vehicle weighing 5 tons which ran at m.p.h. on the Glasgow–Edinburgh Railway.

Little further electrical activity seems to have taken place in cotland until the advent of William Thomson, Lord Kelvin. e entered the University of Glasgow at the age of 10 and carried f many prizes, finally proceeding to Peterhouse College, Camridge. In addition to his academic distinctions, Thomson won e Colquhoun Sculls, founded the University Musical Society nd spent some months in France, where he met many of the ading scientists; in these varied activities he set another example at present engineering students may follow with advantage. homson was thus well equipped to accept the Chair of Natural hilosophy at the University of Glasgow in 1846, where he mained until his retirement in 1899. Professor Dee has spoken the 'Dilemma of Lord Kelvin'—should he devote his abilities pure science or to its engineering applications? As his life rogressed he tended more and more towards the second alternave, one result of which was that he died a rich man. Kelvin ave us the Atlantic telegraph cable, the heat pump, a host of easuring devices and, of course, Kelvin's law; he also invented, oncurrently with Ferranti, the ribbon armature which was used many Ferranti alternators.

Associated with Kelvin in his later years was George Forbes, ppointed in 1873 to the Chair of Natural Philosophy at the nderson College, Glasgow, now the Royal College of Science nd Technology. Forbes definitely abandoned pure for applied cience and set up as a consultant in electrical and civil engineering London. With Lord Kelvin he was invited in 1890 to advise n the first hydro-electric scheme at Niagara and was appointed onsultant to the construction company. It is not very widely nown that, in 1884, he was the first to use the carbon brush.

In these early days two names stand out as the founders of the lectrical manufacturing industry in Scotland—Henry Mavor and avid Bruce Peebles. Henry Mavor entered upon a medical ourse at the University of Glasgow but abandoned that career favour of an association with Col. R. E. Crompton in some of is early electric lighting schemes. Mavor then returned to lasgow and set up arc lighting installations at the Post Office, ueen Street Station and elsewhere; the firm of Muir and Mavor as founded and the first public supply of electricity in Glasgow as given from Miller Street in 1885, so that by that time elecrical engineering had become established in Scotland. Another otable achievement by Mavor was the first application of alterating current to ship propulsion in the 50-ton *Electric Arc*. avid Bruce Peebles, who was born in Dundee in 1866, started firm in Edinburgh to manufacture gas appliances; in 1898, eeing the rising importance of electrical power, he initiated an ectrical department which has now grown into an electrical anufacturing firm known throughout the world for its heavy ectrical plant. Although not born in Scotland, Jens La Cour, ief engineer of the firm from 1903 to 1907, contributed largely its fortunes, and therefore to the fortunes of Scotland, by his ention and development of the motor-convertor.

The name of Munro has been associated with the Scottish ctrical contracting industry since 1840, when David Munro founded, in Glasgow, the firm of Anderson and Munro. In the next generation John M. M. Munro was closely associated with the pioneer work of Crompton and Kelvin in various electrical installations, and in 1888 he built the first electric railway in Scotland from Carstairs House to Carstairs Junction, about 1¼ miles. In 1899 J. M. M. Munro was associated with Kelvin, Henry Mavor and others in founding the Scottish Centre of The Institution.

Scottish engineers have also been represented in the field of communication. Alexander Graham Bell, inventor of the telephone, was born in 1827 and educated in Edinburgh, although he subsequently moved to Canada and the United States, where the actual invention was made. Bell himself in his early years, his father and his grandfather were connected with the teaching of elocution and in vocal physiology, and, as with Watt, the development of the telephone was a spare-time occupation. Another Scottish name in the communication field is that of John Logie Baird, the first man to produce a practical television picture. Like Watt, who got his inspiration when walking over Glasgow Green, Baird is said to have formulated his ideas of television while walking along the cliffs at Hastings, where he had gone in search of sunshine. Baird achieved, perhaps, less than he deserved, as ill-health and lack of finance were often a handicap; nevertheless, starting with very primitive equipment he developed his mechanical scanning apparatus to such perfection as to be able to televise the Derby and other events in 1932. The world was then, however, entering the electronic age and another Scot, Campbell Swinton, as far back as 1908 had suggested that the cathode-ray tube might be applied to the problem. Although Baird's system was used for public transmission by the British Broadcasting Corporation, it could not compete with the alternative electronic system and its use was abandoned. Even after this setback Baird continued with experiments on colour television and other things until his health finally broke down in 1946 at the age of 58.

The last of the Scottish pioneers to be mentioned is Sir Edward McColl. He travelled the hard way—John Brown's shipyard, Glasgow tramways, Pinkston power station, the Clyde Valley Power Co., and finally, in 1927, Manager of the newly-formed Scottish Area of the Central Electricity Board. The McColl protective system became famous in both hemispheres, and long before the Second World War McColl had worked out plans for a pumped storage scheme at Loch Sloy. When the North of Scotland Hydro-Electric Board was formed in 1943, McColl was an obvious choice for the position of General Manager, and the economic success of the scheme, the greatly improved Highland amenities and the revival of local industries resulting from it are fitting memorials to one of Scotland's most famous engineers.

The names referred to are only a selection of those who, in the past, have made Scottish engineering known throughout the world—but what a heritage for so small a country and what a standard to maintain in the future. The students of to-day, and their teachers, have a great task before them if they are to maintain these high traditions.

The purely academic education available to undergraduates in Scotland has long been, and still is, at least equal to that obtainable elsewhere; particular attention is, however, now being paid to full-time and part-time post-graduate courses and to helping the student with such semi-technical matters as public speaking and acquiring an awareness of engineering achievements beyond the limited curricula of his lecture rooms.

Provision of a broad practical training has always been difficult in Scotland as no individual Scottish firm is sufficiently large to give, within its own organization, a training on as wide a scale as that provided by some of the large firms in England. A

training scheme involving co-operation between a number of complementary firms seems, therefore, the only way in which Scotland can make her proper contribution to the practical training of British professional electrical engineers. Such a possibility was mentioned in my Chairman's Address to the South-East Scotland Sub-Centre in 1954 and was subsequently discussed by Sir Hector Hetherington of Glasgow University, Mr. William Fraser and Mr. J. S. Hastie in the West of Scotland, by Mr. W. B. Laing and Mr. C. M. Beckett in the East of Scotland, and also by a number of others, with the result that on the 12th January, 1956, Mr. Fraser convened, at the St. Enoch Hotel, Glasgow, a meeting of representatives of industrialists and academic bodies to discuss the matter. As a consequence a working committee was nominated to investigate the possibilities and prepare a draft scheme for the vacation and post-graduate training of electrical engineering students by making use of the joint facilities of the Scottish electrical industry; a secondary objective was to interest senior school pupils in the possibilities of an engineering career.

After six months of intensive work under the chairmanship of Mr. Kenneth Atchley, the committee produced a scheme which was submitted to the nominating body and, with only small modifications, accepted by them on the 11th July, 1956. This date may therefore be regarded as the birthday of what is now known as the Scottish Electrical Training Scheme (S.E.T.S.), although, since one of the recommendations of the committee was that the scheme should be operated by a non-profit-making company, it was not finally incorporated as such until the 25th February, 1957.

The company comprises at present seven member firms (two heavy plant manufacturers, one switchgear manufacturer, one cable manufacturer, one light plant manufacturer and the two Scottish Electricity Boards; in addition a large electrical contracting firm is about to be admitted); it is managed by a board of seven governing directors representing the member firms and six advisory directors from the universities and leading technical colleges and one from the Scottish Council (Development and Industry). The first chairman of the board is Mr. Atchley.

Shortly after the birthday meeting an executive committee was appointed by the board and began work on the detailed organization. On the 1st January, 1957, a full-time and fully independent organizing secretary was appointed with appropriate staff and office accommodation, and recruitment of trainees commenced forthwith. That this was possible within one year of the original meeting at the St. Enoch Hotel was a notable achievement, and there is no doubt that many future Scottish engineers will owe a considerable debt of gratitude to Mr. Kenneth Atchley for the very great skill and enthusiasm with which he guided the working committee, and later the board, through the many legal, financial and academic difficulties encountered in putting the scheme on a sound basis.

This Scottish scheme offers a two-year graduate course as well as vacation and pre-college training; this summer 59 trainees have been enrolled, including seven graduates embarking on a two-year course, 37 vacation trainees and 15 pre-college trainees. It is intended that about 20 graduate trainees per year will normally leave the scheme as fully trained professional electrical engineers.

Each student spends periods of appropriate length in the organization of one of the member firms, the programme being arranged to give as wide an experience as practicable and a training in accordance with the recommendations of The Institution for practical training of professional engineers. Since Scottish engineering skill is renowned throughout the world, there is no doubt that trainees under the scheme will be in contact with the best modern engineering practice, and particular advantages are the very wide range of experience offered and the fact that this includes not only manufacturing experience but also, through the membership of the Electricity Boards and a contracting firm, experience in the utilization of the manufactured products. An important feature is that trainees are employees of S.E.T.S. and not of one of the member firms.

An obvious problem arising with a joint scheme of this nature is the maintenance of a corporate spirit among the trainees; this is achieved by periodic conferences of all trainees together with senior staff of the member firms, by the regular issue of a bulletin and by the fact that the scheme has premises in Glasgow which can be visited by the trainees and which act as a focal point for the organization. The first conference was held in September and the evidence from it indicates that S.E.T.S. has already become a major factor in the training of Scottish professional engineers.

With regard to the secondary objective investigated by the original working committee, it is proposed to organize short courses of 3 or 4 days each for selected schoolboys during their holidays; each course will comprise about four boys, and, since the activities of S.E.T.S. members cover the whole of Scotland, it is hoped that it will be possible to give opportunities to boys all over the country without an undue amount of travelling or staying away from home.

At present entry to S.E.T.S. is, for various practical reasons, limited to Scottish students or students intending to make a career in Scotland. It is my personal hope, however, that in the future the doors will be opened to students from England and Wales, from the Commonwealth and from foreign countries so that Scotland will be able to play her traditional part not only in training British engineers but in training engineers for service throughout the world.

It will require the combined efforts of Members, Graduates and Students of the Scottish Centre so to formulate the history of the future that a Chairman 100 years from now will be able to continue this great story of Scottish engineers.

The Institution of Electrical Engineers
Abstract No. 2469
Jan. 1958

SOUTH MIDLAND CENTRE: CHAIRMAN'S ADDRESS

By L. L. TOLLEY, B.Sc.(Eng.), Member.

'AUTOMATIC WORKING IN THE TELEPHONE TRUNK NETWORK'

(ABSTRACT *of Address delivered at* BIRMINGHAM, *7th October*, 1957.)

In transmitting communications the objective is to save time ithout impairing the content of the message. Before man arned to write either he had to give the time to go and speak is message or he had to risk the unreliability of a messenger peating a spoken message. A written message is not altered in ansmission and a postman can carry a large number of letters a time. A spoken message recorded on magnetic tape gives n even more reliable transmission, since it carries the inflections f the voice, which are often quite important, but it is necessary have play-back apparatus at the receiving end. A telephone stem carries the spoken message with a very faithful reproduc-on and with an extremely short transmission time. Time can so be saved by those using the system and by those who operate if automatic switching is used to set up the connections. Some 7% of all the telephones in this country are now served from cal automatic exchanges. The introduction of automatic orking on the trunk system is planned, and an automatic cchange for trunk circuits was brought into service in London 1954. A similar exchange is to be opened in Birmingham uring this autumn, and the subject therefore seems very suitable or me to talk about.

Since automatic production and handling are not usually conomic for small quantities, let us first consider the growth of e trunk telephone service as indicated by the records of the umbers of trunk calls:

12 months ended 31st March	*Trunk calls (millions)*
1920	54
1930	117
1940	117
1950	235
1957	321

he appearance of a standstill between 1930 and 1940 is due to e fact that from 1935 onwards some calls have been classified local that were previously classified as trunk calls. Thus the verall increase of trunk traffic from 1920 to the present year as in fact been greater than the figures show. The number of eech channels over 25 miles in length has been increased very onsiderably:

At 31st March	*No. of speech channels*
1920	1 707
1930	3 680
1940	6 220
1950	16 340
1957	22 289

The growth of the service and of the system has both stimulated d been helped by developments in technique. The trunk lines ed at the beginning of the century were of heavy-gauge copper p to 800 lb per mile) carried on pole routes. The introduction inductive loading before the 1914–18 War, and of amplifiers bsequent to it, enabled the conductor weight to be reduced 20–40 lb per mile, a small fraction of what was previously d, and made it practicable to put the long-distance circuits cable. Carrier-current working, which was introduced on erhead lines in the 1920's, was developed for use in cables in 1930's and will now provide 960 circuits on two coaxial s of $\frac{3}{8}$ in diameter. Unless these developments had been eved it would be necessary to have forests of poles supporting many thousands of tons of copper wire to provide the number of circuits now in use—if indeed the circuits would be wanted in such circumstances and at the cost that would be involved.

The trunk system must provide not only for the frequent calls between large centres of population but also for the occasional call from a small town or village to another small place at a distant point. To ensure the efficient use of plant it is obviously necessary to route these occasional calls via collecting centres so that the long-distance circuits between the centres can be used with a good occupancy-time, and for telephone traffic routing the country is therefore divided into Zones which in turn are divided into Groups, one exchange in each Zone or Group being the collecting centre at which the main routes terminate. A call from a subscriber on a minor exchange to a subscriber on another minor exchange in a different part of the country is routed from the originating exchange to the home Group centre, then to the home Zone centre, the distant Zone centre, distant Group centre and so to the required exchange. This is the basic system, but it is by no means rigid: if there is enough traffic between the exchanges in one Group and those in another, a route (i.e. a number of circuits) is provided directly between the two Group centres, and similarly if two individual exchanges have sufficient traffic a direct route is provided between them. However, in general the Group centres collect the trunk traffic and each Group centre has a route to its own Zone centre, and each Zone centre has routes to the others.

Most of the exchanges handling trunk calls at Zone and Group centres are worked manually. Twenty-five years ago all trunk calls were passed forward from operator to operator and the last operator dialled the wanted number (if the called subscriber was on an automatic exchange). It was necessary that the dialling should be done by the last operator because the d.c. pulses which the exchange switches require for their operation are not transmitted satisfactorily over a long-distance amplified circuit designed for use as a speech channel. Towards the end of the 1930's, however, a method of signalling over trunk circuits was introduced using line currents of two frequencies (600 and 750 c/s), the line currents being converted to d.c. pulses at the receiving end. These frequencies are in the speech range, and the method is often called 2VF signalling. The equipment was not suitable for the operation of circuits in tandem, and the system was applied only to circuits between Zone centres, but since the Zone centres are in the larger cities which, having large populations, originate and receive correspondingly large amounts of traffic, the introduction of 2VF signalling did enable the increasing traffic to be handled without so large an increase of operating force as would otherwise have been necessary. A call can now be dialled by the Zone operator to a subscriber on an automatic exchange within dialling range of the distant Zone exchange; e.g. a call originating at an exchange in the west of Scotland can be dialled by the Glasgow operator to a subscriber on an automatic exchange in or near Birmingham, and it is not necessary to call in a Birmingham operator. Similarly, Birmingham operators can dial subscribers' numbers in London, Glasgow, Manchester, Leeds, Bristol and other Zone centres.

Since the 1939–45 War the 2VF equipment has been further developed, and the later form is suitable for use on circuits in tandem. With this development and as automatic trunk exchanges are installed it will be possible for the operator at

Mr. Tolley is in the Post Office Engineering Department.

the originating end to set up a trunk call by dialling through the trunk exchanges *en route* and also to dial the wanted subscriber's number if the target exchange is automatic. Consider, for example, a call from Holmes Chapel in Cheshire to Whitstable in Kent when the Birmingham automatic trunk exchange is in service. The originating subscriber will dial 'O' and (there is no operator at Holmes Chapel) the equipment will automatically connect to the switchboard at Crewe: the Crewe operator will dial through the Birmingham trunk exchange for London, through the London trunk exchange for Canterbury, through Canterbury exchange for Whitstable, and will then dial the wanted number. Only one operator will have been required to set up the call. As automatic exchanges are opened at other Zone centres and the necessary equipment is installed at Group centres, the range of dialling over the trunk system will be increased until all exchanges can be reached in this way. The total amount of plant to be installed is very considerable. The installation at Birmingham is costing over a million pounds, and even so the forecast of traffic indicates that the exchange will have to be enlarged in a few years' time. Although the installations in Group centres are smaller than those in Zone centres, there are, of course, many more Groups than Zones.

The next stage in converting the system to fully automatic working will be to make it unnecessary for even one operator to come into circuit but to provide for the caller to dial his call right through. I must assume, of course, that both the terminal exchanges are automatic and that the necessary equipment has been installed at the intermediate Zone and Group centres. There will, however, be required some additional equipment at the originating end to route the call and to measure the charge.

To explain the routing requirement let us consider again the example of a call from Holmes Chapel to Whitstable. When the operator dialled the call she (or he) needed to know what to dial in order to position a Birmingham switch on a London circuit, then to position a London switch on a Canterbury circuit, then a Canterbury switch on a Whitstable circuit. This information is given in files which show the routings for calls, and such files are provided for operators at all exchanges. The files differ from one exchange to another because the routing from, say, Bournemouth to Canterbury is different from that from Crewe to Canterbury, but an operator serving in a particular exchange needs only the file for that exchange. If subscribers who are to dial their own trunk calls were to use similar routing files they would need, not only the file for their home exchange, but also files for other places that they might visit; and additions would have to be made from time to time as new automatic trunk exchanges were opened. The arrangements will be very much simplified by giving a number to each exchange in the country so that the subscriber may dial the number of the exchange that he wants followed by the number of the subscriber on that exchange. Equipment will have to be installed to register the digits dialled and to translate those digits that constitute the number of the required exchange into whatever routing digits may be required to take up circuits and establish a connection to that exchange. The function of this register-translator will be very similar to that of the equipment now used in the larger cities to receive the three 'letters' of the exchange name and route the call. The equipment for trunk dialling will also need to store the digits of the wanted subscriber's number, as the present equipment in the large cities does, until the connection has been set up to the wanted exchange and will then repeat them into the distant exchange equipment.

To establish the charge for a trunk call that is set up manually the operator makes out a ticket showing the calling subscriber's number and the details of the call. The charge to be made depends on the distance between the terminal exchanges and on the time that the call lasts. It would be practicable to construct apparatus to produce similar tickets automatically, but this would be complicated and costly. The register-translator, which, as I have already mentioned, will set up the routing of the call, must receive the digits dialled, and as the first few digits will indicate the wanted exchange (and thereby the distance), this equipment can determine the rate for the call. When it has routed the call, has sent forward the remaining digits and has signalled to recording apparatus the rate of charge, the register-translator can be released and become available for another call. The recording apparatus must remain in circuit to record the total charge, which depends on the time the call lasts.

Before subscribers on a given exchange can dial their own trunk calls the automatic switchgear must be provided at the trunk centre and the register-translators must be installed; and for reasons of manufacturing capacity, installation capacity and expense it will be possible to make the facility available only by a gradual process. Equally, the number of exchanges that can be dialled will be limited at first and will increase as further installations are completed. The change must necessarily be spread over a number of years, and to make the whole system automatic we must also convert to automatic working the 1 200 local exchanges that are at present manual. Although operators will not then be required for setting up trunk calls, the operator will not disappear from the telephone organization. He (or, more generally, she) will be required for assistance and inquiry calls, and it should be remembered that the amount of this traffic will be related to a much increased total of trunk traffic. In fact, if all the trunk traffic expected in the future had to be set up manually the operating force required would represent a serious drain on the national resources. In the same way that the growth up to now could hardly have been achieved on open-wire routes and has become a practical possibility by reason of the development of the cable routes, so the growth expected in the future could hardly be achieved if operating were to be continued on a manual basis.

Papers have already been read and others will certainly be read in future to describe the methods and the equipment used in these developments, and I have given only an overall review of what has been done and what is intended. It will be clear that the introduction of automatic working in the telephone trunk network results from and requires the use of a number of techniques; particularly it requires the co-operation of transmission engineers and switching engineers. Co-operation between men engaged in different fields of activity is indeed a necessity in every large project. The whole field of engineering development is very wide, for many years three major engineering Institutions have dealt with different parts of it, and within our own Institution Specialized Sections have been set up as electrical engineering has become more diversified. It is important that specialization should be provided for and encouraged, since to achieve the major advances in knowledge, design and development we need engineers who specialize closely. It is also important to ensure co-operation and to provide for co-ordination in the larger projects, and therefore the industry also needs men to appreciate the problems in several fields. We must all specialize to some degree, whether more closely or less, but as the Institution programmes both of the Centres and of the Groups are varied within their fields, they do enable us to enlarge the area of common ground that we share with others, provided that we do not limit our attention only to those meetings dealing with the work on which we are already engaged. It is, of course, essential that there should be some common ground from which co-operation can start, and as we all know it is essential that we should co-operate in order to get the maximum development.

The Institution of Electrical Engineers
Abstract No. 2506
Jan. 1958
©

SOUTHERN CENTRE: CHAIRMAN'S ADDRESS

By L. G. A. SIMS, D.Sc., Ph.D., Member.

'NOT WITHOUT HONOUR—A CONTEMPLATION OF UNIVERSITY, COLLEGE OF TECHNOLOGY AND STUDENT MEMBERS'

(ABSTRACT *of Address delivered at* PORTSMOUTH, *2nd October*, 1957.)

I should like to use for my Address the theme of engineering ucation and to review the functions of university and technical llege as they appear after a period of renewed evolution. There e now the new Colleges of Advanced Technology, with their -month sandwich diploma courses, and the new Regional olleges of Technology. These represent changes in the picture higher education and perhaps call for some re-appraisal of sition by university applied science departments.

In speculating upon future university developments, I must ake it clear that no official university policy is implied by my ggestions.

A successful approach to communication, to the teaching of idents, is to work almost exclusively in contact with them. his makes the teacher an artist in the interpretation of syllabuses d regulations, both of his own college and of the many ternal examining bodies. My thoughts are with numerous st and present colleagues whose teaching abilities justify that scription. Some of the colleges whose routine is served best this system have now been promoted to overcome the nation's esent shortage of technologists. These are the new Colleges Advanced Technology. There are also the newly-termed egional Colleges of Technology, and finally there are the neral technical colleges, some of which are large, with modern ildings, and with good staffs and equipment.

In all these colleges the emphasis is upon teaching, though t to the exclusion of research. The banner of their pride wears e legend: 'This communication of knowledge is important.'

In the universities the authorities place emphasis upon sidence and research. This means both experimental research d research which is concerned with reading and the exercise the mind. These forms are needed if some university students e to be inspired by professors, lecturers and tutors to ascend the highest peaks of thought, to discern new lands and to ite their names later as discoverers in the long story of man's hievement. We think, therefore, that every university engineer- acher should spend appreciable time upon scholarly pursuits, cluding study abroad, and that he should be provided with odern equipment and with extensive library facilities. His work ould be 'predominantly intellectual and varied'. The dis- guished academic colleagues with whom I have worked, and ll have the pleasure to work, would endorse that opinion. An mosphere is provided thereby which lies about the student and otly enables him to unfold, develop and flourish.

University academic staff exercise the privilege of entrusting earch projects to the hands of their graduates, sometimes der Government contract, and of recommending those aduates, when successful, for research degrees (also called gher degrees).

In the university itself and in its halls of residence, engineer- aduates and undergraduates enjoy the company and ideas of idents from other faculties, such as those of science, arts, eology, economics and law. But student social life to-day is exclusive to universities than was formerly the case, due to growing residential aspect of other colleges, and indeed not only to that. Most large engineering firms have taken over fine old country houses as residences for their various grades of apprentice. Names such as Coombe Abbey, Coton House and Castle Bromwich Hall come to mind. In fact the engineering industry is housing its young people not only in hostels and halls of residence but in the stately homes of England.

Sandwich Courses in Electrical Engineering.—There is now a form of entrance to Southampton University and its electrical engineering courses, called the thick sandwich or 1–3–1 scheme, in which combined studies are made available by the University and various branches of industry. It combines academic educa- tion with industrial training. One superimposes this scheme upon the University's traditional pattern of education for electrical engineers by giving the sandwich a thick filling which is in fact the full-time three-year degree course. We adopt and adapt earlier systems (consisting, for instance, of a vacation year in industry between school and the University) and make of them a complete five-year course of education and training. Variations of the 1–3–1 scheme are possible, without disturbing funda- mentally the timetable routine of the University. The training curricula of these sandwich courses are at present wholly in the hands of our industrial friends.

This development in combined university education and electrical works-training is now on trial, and it is early to judge its results against those of the direct-entry system. We expect that the latter will always be required, because enlightened industry looks for a certain number of young graduates who will challenge current works procedure and desire to introduce new ideas. That is presumably most likely to happen with a few of the more gifted direct-entry graduates, because they will not have experienced any previous industrial influence.

On the other hand, I have heard a leader in industry state fairly recently, and repeat, that he prefers young men who are primarily works apprentices, who have gained their academic en- gineering education concurrently by some form of part-time study.

It is assumed that university sandwich courses improve upon that point of view, even though not complying literally with it.

University and industry at present fix their own curricula for such courses independently, the one in education, the other in training, but it may prove advantageous, at some future time, to introduce a measure of joint consultation. In the meantime, industry gives encouragement to the university sandwich experi- ments and has proved eager to help, whenever and wherever consulted. Joint consultation between university and industry could open wider the door to co-operation between firms, in the national interest, and to inter-firm student training, where that is desirable. Southampton University, on its part, already shows its keen interest in the students during their works train- ing, by tutorial visits.

We have also a means of entry for the best examinees amongst young men after success in H.M. Dockyard Schools examina- tions or after their Ordinary National Certificate year. It is called 'industrial entry'. It meets the needs of those who have not come solely by 'A' level results in the G.C.E. examinations.

The Arts Graduate in Industry.—It is important that the re- turns for 1955–56 of the University Grants Committee (if a recent

r Sims is Head of the Electrical Engineering Department, University of th .mpton.

statement about them has been interpreted correctly) show 43% of all university students as occupied in the arts, against 21% in the pure sciences and only 13% in the technologies.[1] Clearly there is room for a point of view in which industry employs arts graduates.

The picture of a place in industry for the arts graduate is illuminated by information supplied by a works education and training department, about some of their arts recruits, who, it is stated, start work in the training drawing office and training workshop. After the first six months they go into departments dealing with organization of various types, but factory departments are not likely to form a very significant part of the training. They have the opportunity of studying one day a week to give them some scientific background, 'but there is no intention whatever of turning them into engineers'.

Nevertheless, it could hardly be regarded with equanimity if young non-engineering people entering engineering firms and not carrying any direct engineering responsibility should gravitate —or levitate—into many of the senior administrative posts. There is perhaps a tendency to regard an arts education as suiting a man for government and a science education as fitting the other man for the back room only.[2]

History and Philosophy of Science in Engineering Courses.—Perhaps one may refer briefly to the inclusion of some study of the philosophy of life and nature in engineering teaching.

The best time to inculcate the idea of a 'wisdom of the past' in the mind of the young engineer is during his schooldays, and a broad thread of historic and humanistic outlook should continue thereafter during his university career. It is for that reason that a close tie with classical physics as well as with modern physics is desirable at the university. This can be accomplished by introducing a special subject, called, for instance, 'history of science'.[3,4] But experience makes me doubt whether it is attractive to students and, if not, whether it can be wholly successful. On the other hand, a well-conceived course of study in the broad field of physical-engineering measurement provides an opportunity of discussing the pioneer work in physics and engineering and the personalities of the pioneers, many of whom were great characters. In addition, it continues the fundamental electricity and magnetism and is appropriate to all electrical engineering students. Good students re-live in their own laboratories some of the difficulties and triumphs which men like Kelvin, Clerk Maxwell, Latimer Clark, and other intellectual scientists of the nineteenth century experienced when establishing the theory and practice of accurate electrical measurement. Precise measurement in engineering is remarkably stimulating and self-rewarding to students and represents in their minds an approach to ultimate perfection.

Longer Engineering Courses.—In face of the advancing front of scientific knowledge the academician to-day finds it a great question how much he should attempt to impart to engineering students in a three-year university course. It may well be that the limits of the present three-year courses have already been reached, although a device exists, a palliative, used in the engineering courses at Southampton and elsewhere, namely that of dividing the engineering course into fast and slow courses at the end of a common first year. The fast course reaches honours standard in three years and the slow course is of the same duration but has less academic content and is for students of average ability. It leads to an ordinary pass degree and has an advantage in that more time is available for class exercises and examples.

In electrical engineering it is at least debatable whether the fast honours course, though well-conceived originally, still covers sufficient ground to prepare university honours students for the more recent forms of modern industry.

Another palliative is to emphasize certain electrical subjects more than others, by subdividing the honours men into two streams, for example those who select power studies and those who prefer electronics.

Now subdivision of electrical engineering instruction should only be carried a short distance, at least in a university of moderate size, and I should therefore like to make the perhaps unexpected suggestion that still larger first-year university engineering classes should be contemplated.

At the end of the first university year, it would be assumed that examination results then gave a true picture of the developed schoolboy, and the universities could divide the first-year group into two or more sections, selected upon examination marks. It would be easy to say that this procedure would divide the men into sheep and goats, but that phrase is no longer considered to be realistic. The goats are no longer unutterable, but represent valuable junior engineer personnel whose abilities are matched to some other form of engineering education, perhaps to the three-subject curricula of the National Certificate schemes, beginning perhaps at the A.1 stage and finishing with a Higher National Certificate. The further goal of Part III of the new Institution Examination, with Associate Membership to follow, should be regarded as attainable by the best of these students. This principle of dealing with the lower-half groups amounts to limiting the range of studies but maintaining a high standard.

We have considered 'fast' and 'slow' courses. The logic of developments, having regard to the present state of scientific knowledge, is to consider 'long' courses. These could be basically of two forms, the first an extension of the present three-year honours course, through what is now known as post-graduate study. For these developments, some temporary industrial restraint in recruitment and some industrial finance might be needed.

British long courses of four, eventually of four and a half or even five years' duration, could excel the engineer-diploma courses of the Continent. But only those students who proved themselves able enough academically to benefit from the long courses fully would be drafted into such advanced studies.

The Master of Science Degree.—The custom in British universities hitherto has been to set the three-year graduate or bachelor, who is seeking a year's further study, an experimental problem in which both theory and experiment are involved. This Master of Science course includes staff supervision but probably no lectures, over a period of at least one year, but latterly the inducements offered by industry have taken away most young graduates at the first-degree stage: in many cases this is to be deplored. A revival and some further development in the M.Sc. form of study for fourth-year engineer students is desirable. For many university students it is still a good finish to their academic work and it does not keep them too long from industry.

But new 'long' courses reaching the present post-graduate theoretical levels of the largest engineering schools should also become a regular part of all university engineering education.

References

(1) *The Times Educational Supplement*, 26th July, 1957, p. 1054.
(2) Ratcliffe, J. A.: 'Science as Part of a General Education', *New Scientist*, 4th July, 1957.
(3) Collman, R. V.: 'The Value of History', *Journal of the Engineers' Guild*, 1956, **4**, p. 329.
(4) Hanson, N. R.: 'The History and Philosophy of Science in an Undergraduate Physics Course' (Institute of Physics reprint).

The Institution of Electrical Engineers
Abstract No. 2497
Jan. 1958
©

EAST-ANGLIAN SUB-CENTRE: CHAIRMAN'S ADDRESS

By G. E. MIDDLETON, M.A., Member.

'EDUCATION FOR ENGINEERS'

(ABSTRACT *of Address delivered at* NORWICH, *7th October*, 1957.)

My justification for an address on this hackneyed subject is that a teacher I enjoy confidences from pupils, parents and indus-ialists which are often very revealing. By engineers I mean hartered Engineers, and of their importance in to-day's society ere is no doubt, but perhaps because their standards are deter-ined by reference to facts and not to opinions, these engineers e not always properly understood as persons in social life. The ngineer's scientific methods give him confidence in his own apabilities, and this sometimes leads to arrogance but more often an understanding humility.

Education I feel has three broad aims: to teach understanding life as a whole; to foster genius; and to make a man useful. o achieve understanding is a very personal task, and many ise men have given us glimpses of greatness throughout the ges. I would only emphasize here that true education must rovide ready access to books of all kinds, and, particularly for ngineers from their earliest years, access to men of affairs. ngineering is a live activity, and its special ideas are better eveloped by active interchanges between people. Older engi-eers should go out of their way to reach intimate understanding ith the younger, so passing on some of their philosophies at rst hand to the developing generation. Education is for the nprovement of individual capabilities, and the encouragement genuine interests, not the imparting of set ideas through termediaries.

To foster genius is vital for our country's future; my belief is at our finest national contribution to affairs is originality. We nnot, then, copy foreign methods, however streamlined and ficient they may be. When a teacher recognizes even a little enius in his pupil he should refrain from inflicting routine tasks nd unnecessary exercises, and he should actively encourage his upil's special bents. Academic discipline should be used only guide zestful energy, not be made an excuse for quotas of ritten work.

An engineer, by definition, is a useful man. He thus has to arn words and techniques in a wide range so that he can deal ith likely situations. His education must include a wealth of tablished material, but this should not overwhelm his teaching. ormal technical education is well provided for in this country, nd we have enthusiastic teachers in all industrial areas. Most olleges I have come across have, however, proclaimed syllabuses hich are far too ambitious for most pupils, and for them learning too much of stark fact and bare routine. The pupils appear not have time to do any critical thinking, nor are they encouraged do any for examination purposes. Colleges could well review eir work, trying less stereotyped questions in their papers, and aching—even re-teaching—more examples of sound principles an of direct applications, especially in the courses for Higher ational Certificates.

It is, however, at graduate level that so much more could be one by industry. Graduates are already trained for quicker opreciation of ideas than ordinary apprentices, yet so many are ustrated by slow and unimaginative time-tables. In a paper[1] esented to The Institution of Mechanical Engineers recently, e authors describe techniques which were so successfully plied to groups of undergraduates in factory courses that very eful and stimulating results were obtained. The enthusiasm nerated was most striking. The same paper contains a searching analysis of the special attributes of graduates, and this should provide food for thought in those responsible for their training. Too many firms appear to waste bright men—and not only apprentices—in dull and discouraging tasks, and much could be done to relieve the shortage of technologists by encouraging and trusting the young men in responsibility.

One excellent way of developing a young engineer is by providing tutors in industry who can offer guidance and experience within the live environment of manufacture and research. Too many of our best students spend extra years at a university carrying out restricted research investigations which lack the life or urgency of the corresponding industrial problems. The university's function is to provide a three years' basis for future mental development, but more time than this is often devitalizing for the engineer. Where industry is prepared to be generous with tutorial time, the young graduates can tackle responsible problems and obtain useful answers in cases whose unusual nature makes them an embarrassment to the time of the regular staff of busy men. In Cambridge we have first-hand knowledge of many firms in which apprentices under guidance have conducted urgent technical investigations very successfully, and so have obtained a much deeper experience of practical realities than formalized training would give. Money spent on such tutorial effort is to my mind better spent than on increased stipends for teachers of young students who are without any knowledge of realities or of the complexity of real situations. An extension of this tutorial effort could be made to provide some new degree, or such distinction, for high-grade intellectual work carried out in the industrial context. This would offset the temptation now presented to young men to spend too long a period in the soft atmosphere of a teaching establishment purely to obtain a higher degree in the hope of impressing a prospective employer. Only a few men have the imagination which produces its results in academic isolation; most brilliant engineers have minds which are stimulated and developed best by fast-moving, vital and creative work in a good firm.

I have been reading a recent biography[2] of Brunel, the engineer. His father was able to give him the finest education available in the 1820's, and young Brunel was accordingly taken from school at the age of 14 to go to what might be termed a high school. By the time he was 16 he had completed his schooling and a good apprenticeship as well. He became engineer-in-charge of construction of the first Thames tunnel at the age of 20, and he was coping with a succession of crises of the sort which would daunt many a modern engineer. Where, in all the tidy and over-comprehensive schemes of to-day, could we discover and foster a man of such abilities? Let our teachers and tutors beware of losing the diamonds in the mass-flow curricula of our schools. Let us not forget that in giving too much instruction we stifle originality, elegance and fun, and all these factors are vital to man if he is to be worth while.

Mr. Middleton is a Lecturer in Engineering at the University of Cambridge.

REFERENCES

(1) MARPLES, D. L., RADFORD, J. F. A., and REDDAWAY, J. L.: 'An Approach to the Techniques of Graduate Training', *Proceedings of The Institution of Mechanical Engineers*, 1956, **170**, No. 22, p. 747.

(2) ROLT, L. R. C.: 'Isambard Kingdom Brunel—a Biography' (Longmans, Green and Co.).

621.396.11

The Institution of Electrical Engineers
Paper No. 2427 R
Jan. 1958

THE ATTENUATION OF RADIO WAVES REFLECTED FROM THE E-REGION OF THE IONOSPHERE

By R. W. MEADOWS, B.Sc.(Eng.), Associate Member.

(*The paper was first received* 1*st January, and in revised form* 12*th July*, 1957.)

SUMMARY

Measurements of the absorption of waves travelling between Slough and Inverness (740 km), and reflected once from the E-region during the process, are compared with similar measurements made simultaneously for vertical incidence at Slough. The absorption over the oblique path, calculated by Martyn's absorption theorem from the value obtained at vertical incidence, was found to be much too low. Approximately correct values are obtained from the formula for non-deviative absorption, provided that the lowest-frequency (2 Mc/s) vertical-incidence results are used: otherwise the oblique-incidence value is too high. It is therefore considered tentatively that the absorption measurements which have been made for many years past at 2 Mc/s, using vertical-incidence reflections at Slough, might prove to be a useful guide to the absorption likely to be obtained over other oblique E-region paths during any part of the sunspot cycle, and they have been formulated accordingly. The anomalous variation of absorption which occurs during winter in high latitudes is also considered.

The additional transmitter power required, as a consequence of fading, to enable a desired signal level at the receiver input to be exceeded for given percentages of the time is discussed, and a method of calculating losses due to polarization changes brought about by magneto-ionic splitting is indicated.

(1) INTRODUCTION

The paper is intended primarily as a contribution to the problem of calculating the attenuation likely to be obtained over ionospheric paths from absorption measurements made at vertical and oblique incidence, with particular reference to one-hop E-region trajectories at frequencies of, say, 1·5–15 Mc/s. Some new experimental results of absorption measurements are incorporated in the paper, and these, together with absorption data already existing, are expressed in tentative empirical formulae to enable estimates of field strength and fading to be made for short paths.

Measurements of absorption on short paths have been made before by, for example, Piggott and Beynon,[1] and Allcock,[2] all of whom concentrated primarily on F-region trajectories. The series of measurements incorporated in the present paper, however, involves E-region trajectories. This has two advantages: the possibility of losses due to partial reflections[1,3] is remote, and the obliquity obtainable is greater since the height of reflection is lower. Hence there seemed a better prospect of testing the validity of Martyn's absorption theorem (which expresses oblique in terms of vertical-incidence absorption) on a given short path than would be obtained by using F-region measurements. Special measurements of the amplitude of pulses reflected from the E-region were therefore made at 5·1 Mc/s on a 740 km north-south path between Slough and Inverness, and at 1·62 Mc/s at vertical incidence at Slough, the latter frequency being near the Martyn equivalent value for the oblique path. Noon values of absorption at vertical incidence, measured daily at Slough at a number of frequencies, are also examined from the point of view of testing the accuracy of the commonly-used formula for non-deviative absorption when applied to the oblique path.

It is well known that the great variations due to fading limit the accuracy with which measurements of absorption can be made, but little appears to have been published in the past giving details of the errors to be expected, so that difficulty is often experienced in assessing the validity of theories of comparison which are based on the measurements. A semi-empirical process for estimating errors due to fading has therefore been evolved and applied to the present data in another paper.[4] In that paper fading is considered to consist of rapid and slow components, since there appears to be both practical and theoretical justification for this step.

The various factors required for calculating the attenuation in short ionospheric paths will now be discussed under the following headings: spatial attenuation and losses due to magneto-ionic ellipsing, absorption, and fading.

(2) SPATIAL ATTENUATION AND LOSSES DUE TO MAGNETO-IONIC ELLIPSING

Spatial attenuation is the component of path attenuation due to the expansion of the wavefront which occurs as the distance from the source is increased. For short ionospheric paths, when focusing due to the earth's curvature is not great, this results in a power intensity in the wave which varies inversely as the square of the distance from the transmitter. The following approximate expression for spatial attenuation between matched aerials at the ends of a path of length d/λ wavelengths is then obtained:

$$A_s = 60\frac{(d/\lambda)^2}{G_tG_r} \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

This formula gives the ratio of power transmitted to maximum power received, where G_t and G_r are the power gains of the transmitting and receiving aerials, respectively, with respect to that of a half-wave dipole in free space, obtained in the appropriate directions after allowing for earth reflections.

When an electromagnetic wave enters the ionosphere it is resolved into two elliptically polarized components, the ordinary and extraordinary waves, which, for present purposes, may be regarded as being propagated independently through the medium. The extraordinary component is absorbed more than the ordinary at frequencies for which both are propagated, and so the transmitting aerial should ideally be designed to excite the ordinary component only, and the receiving aerial should couple in the optimum manner with the ordinary-wave ellipse incident upon it. However, this can rarely be done in practice, and losses then occur which can be estimated by the following general procedure.

The radiation from the transmitting aerial will be assumed to be linearly (plane) polarized, and of intensity E. It is possible to resolve the vector $\bar{E}$ into a pair of contra-rotating elliptically

Written contributions on papers published without being read at meetings are invited for consideration with a view to publication.

The paper is an official communication from the Radio Research Station, Department of Scientific and Industrial Research.

larized vectors contained within the wavefront, of equal centricity but with major axes E_{oa} and E_{ea} at right angles, cording to the formulae

$$E_{oa} = E\sqrt{(\cos^2\theta + R^2\sin^2\theta)/(1+R^2)} \quad . \quad . \quad (2)$$

$$E_{ea} = E\sqrt{(\sin^2\theta + R^2\cos^2\theta)/(1+R^2)} \quad . \quad . \quad (3)$$

ere R is any desired minor/major axis ratio, and θ is any sired angle between the directions $\bar{E}_{oa}$ and $\bar{E}$, subject to the ervation that E_{oa} and E lie in the plane of the wavefront.

If now θ is made the angle between the plane-polarized ctor $\bar{E}$ and the component of the earth's magnetic field lying the plane of the wavefront at the base of the ionosphere, and satisfies the approximate magneto-ionic condition[5]

$$R = \sqrt{(1 + \tfrac{1}{2}\tau\sin\epsilon\tan\epsilon)} - \tfrac{1}{2}\tau\sin\epsilon\tan\epsilon \quad . \quad . \quad (4)$$

e two ellipses become, respectively, the ordinary and extra-dinary components corresponding to the limiting polarization[5] the base of the ionosphere. τ is the ratio of gyro-frequency wave frequency, and ϵ is the angle between the ray direction d the earth's magnetic field.

E_{oa} and E_{ea} are then, respectively, the amplitudes of the major es of the ordinary and extraordinary ellipses, as excited by the rticular transmitting aerials used. The major axis of the dinary ellipse lies in the plane containing the directions of the rth's magnetic field and the ray.

The ordinary- and extraordinary-wave ellipses are then propa-ted independently through the ionosphere, the eccentricity of ch changing as the inclination of the ray path to the direction the earth's field alters. However, the energy content of each ipse, after allowing for losses due to absorption and spatial enuation, must remain constant, which means that the size d shape of each ellipse can alter only in such a way that the n of the squares of the major and minor axes remains fixed. the minor/major axis ratio on leaving the ionosphere is termined by a further application of eqn. (4), the final size d shape of the (unattenuated) field ellipses arriving at the ceiving aerial can then be calculated, the major axis of the dinary-wave ellipse again lying in the plane containing the ection of the earth's magnetic field and the downgoing ray. e response of the receiving aerial to the ordinary (or extra-dinary) wave incident upon it can then also be calculated, and mpared with the value which would have been obtained had aerial been designed to pick up all the energy available.

The foregoing will now be applied to E-region transmission m Slough to Inverness at the frequency of 5·1 Mc/s. This th is practically north-south, and will be taken as lying entirely thin the magnetic meridian; consequently the ordinary-ray pse on the Slough side will be polarized with the major s E_{oa} in the vertical plane. A vertical aerial was used at ough, so that $\theta = 0$. Other data for the wave incident on Slough side are: $\tau = 0\cdot26$, $\epsilon = 85°$, $R = 0\cdot33$.

Since $\bar{E}_{oa}$ is in the vertical plane with $R = 0\cdot33$, the ordinary is predominantly vertically polarized; had ϵ been 90° it would ve been entirely so. From eqns. (2) and (3), $E_{oa} = 0\cdot9E$, d $E_{ea} = 0\cdot3E$.

For the wave leaving the ionosphere on the Inverness side, $= 0\cdot26$, $\epsilon = 53°$ and $R = 0\cdot87$. The vertical and horizontal mponents of the (unattenuated) ordinary ray arriving at erness are then $0\cdot72E$ and $0\cdot62E$, respectively, the corre-nding extraordinary components being $0\cdot21E$ and $0\cdot24E$. horizontally polarized receiving aerial was used at Inverness, that the voltage induced into it by the ordinary ray would be 62 times the value calculated from the free-space attenuation mula (1). This represents a loss of 4 dB.

The gains G_t and G_r for the aerials used were estimated to be $-0\cdot4$ dB and -1 dB, respectively, for the appropriate angle of incidence. This can usually be calculated with sufficient accuracy by assuming specular reflection to occur at a height of 100 km for reflections from the E-region. The component of path attenuation due to spatial attenuation and polarization changes is then 105 dB. It is interesting to note that, if a horizontal transmitting aerial had been used at Slough, an additional loss due to polarization changes of at least 10 dB would have been incurred.

If this calculation were performed for the reverse direction, the same result for the transmission loss of each magneto-ionic component would be obtained; the ordinary wave downcoming to Slough would then, of course, be predominantly vertically polarized.

(3) ABSORPTION

(3.1) Some Measurements at Oblique and Vertical Incidence

Measurements of the relative field strength of one-hop E-region reflections obtained at Slough from a pulse transmitter in Inverness operating at 5·1 Mc/s have been analysed in Reference 4. Measurements were also obtained for the reverse direction. The absorption at noon calculated from the ratio of noon to night-time* r.m.s. field-strength measurements was estimated to be about 29 dB, with a 99% probability of this estimate lying within limits of 25 and 40 dB. These limits are asymmetrical because the analysis implies a fading echo which obeys a Rice distribution rather than a log-normal one. The asymmetry becomes less pronounced, however, as the confidence limits are lowered below 99%.

The diurnal variation of the mean absorption is assumed to obey the usual non-deviative law

$$A_i = A\sec i\cos^n \chi/(f + f_L)^2 \text{ decibels} \quad . \quad . \quad . \quad (5)$$

where

A_i = Absorption at incidence angle i, dB.
A = Absorption at vertical incidence for $f + f_L = 1$, dB × (Mc/s)^2.
f = Wave frequency, Mc/s.
f_L = Gyro-frequency corresponding to the (longitudinal) component of the earth's magnetic field resolved along the ray path.
χ = Solar zenith angle.
i = Angle of incidence of the trajectory at the base of the ionosphere, and is the complement of the angle of elevation at the ground for a flat earth.

A plot of log A_i versus log cos χ produces a straight line of slope n, which was found to be in the vicinity of 0·6 for the data used. The constant A was found to be 310 with $i = 16°$. It should be remarked that these values relate to the smoothed diurnal curve in Fig. 4 of Reference 4. A different method of smoothing could lead to different values, especially of the index n. Fortunately, variation of the latter is not of much practical importance, since its greatest effect is at large values of χ, where the absorption is least.

The transmission loss measured for the path Inverness–Slough was estimated to be 136 dB, and 140 dB for the reverse direction between the same aerials. Bearing in mind the limitation in accuracy brought about by the variation due to fading and by instrumental errors (estimated at ±5 dB), the 4 dB difference between these two values is not significant; the paths cannot be said to have been non-reciprocal to this extent. A mean value of 138 dB will therefore be adopted.

As the unabsorbed value calculated in Section 2 was found to

* The unabsorbed field strength at night is difficult to estimate, but see the remarks in Section 5 of Reference 4.

be 105 dB, the absorption at noon implied by the absolute measurements of transmission loss is 33 dB. This compared favourably with the value of 29 dB obtained from relative absorption measurements.

Simultaneous measurements were obtained for vertical incidence at Slough at a frequency of 1·62 Mc/s, which is in the vicinity of 1·4 Mc/s, the frequency equivalent for one-hop E-reflection to the value of 5·1 Mc/s used on the Slough–Inverness path. These results are also analysed in Reference 4. A statistical analysis of the amplitude variation (fading) showed that there was a 99% probability of the absorption for noon over the period considered lying between limits of 30 and 46 dB, with a probable value near 34 dB.

The diurnal variation was examined in the same way as for oblique incidence and it was found that $A = 290$ and $n = 0{\cdot}6$.

(3.2) Comparison between the Results for Oblique and Vertical Incidence

It is conventional to calculate the absorption to be expected on an oblique path from results measured for vertical incidence by applying either the non-deviative formula (5) or Martyn's absorption theorem. The accuracy of these methods will now be tested by applying them to the measurements described in Section 3.1.

The validity of eqn. (5), which assumes all absorption to be non-deviative in character, is discussed in Reference 6, where magneto-ionic calculations indicate that deviative absorption cannot in general be neglected compared with the non-deviative, even at low frequencies, the effect being that the absorption at oblique incidence calculated from vertical-incidence values by eqn. (5) tends to be too high. This means that the value of A in eqn. (5) calculated from oblique-incidence measurements should be lower for the same ionospheric conditions than that calculated from vertical-incidence measurements. However, this is not so in the present case, since the values are 310 and 290 dB × (Mc/s)^2 for oblique and vertical incidence, respectively. This difference is not, however, significant, bearing in mind the spread of the 99% confidence limits assigned to the absorption figures from which the values of A were calculated, and corresponds to a difference of 2·5 dB at vertical incidence at the frequency of 1·62 Mc/s. Evidently the measurements described in the present paper cannot be regarded as testing the accuracy of eqn. (5), since variability due to fading is too great. Oblique-incidence measurements made at lower frequencies on the same path, or measurements on longer paths, would be needed to settle this point.

It is, however, instructive to discuss the routine vertical-incidence absorption measurements, made at Slough at noon daily, in this context. The frequency range covered by these measurements is 2–4·8 Mc/s. Monthly means of noon values of A at each frequency for May, June and July, 1954, covering the period of the measurements described in this Section, were taken. They are given in Table 1 below, together with the values of absorption calculated from them for the Slough–Inverness path using eqn. (5).

The standard correction[7] allows for the deviative absorption introduced in the vicinity of the critical frequency of the E-region, around 3·2 Mc/s in this case. It was calculated assuming a parabolic layer with no magnetic field due to the earth.

Except for the 2·0 Mc/s case, and possibly for 2·4 Mc/s also, it is quite clear that the values of absorption calculated for the oblique path are appreciably too high. It should be pointed out here that the overall A figures published[8] from Slough data are weighted in favour of the lower-frequency measurements, so that the discrepancy when using them is not likely to be as great as that suggested by Table 1; for instance, the values published for June and July, 1954, were 350 and 335, respectively.

Table 1

Absorption on Slough–Inverness Path to be expected from applying the Non-Deviative Absorption Formula to Vertical-Incidence Measurements (1954)

Frequency	Noon absorption (vertical incidence)				A	Standard correction to A	Absorption calculated for 5·1 Mc/s Slough–Inverness
	May	June	July	Mean			
Mc/s	dB	dB	dB	dB	dB × (Mc/s)^2	dB × (Mc/s)^2	dB
2·0	29·8	28·4	30·1	29·4	322	0	34·1
2·4	25·2	29·0	29·1	27·8	392	−20	39·6
2·8	25·9	30·4	24·6	26·1	466	−50	43·3
3·2	28·6	28·7	27·8	28·3	600	−∞	—
4·0	17·3	17·8	21·0	18·7	550	−60	52·0
4·8	15·1	18·1	16·3	16·4	646	−70	61·2

It appears from the theoretical calculations of Reference 6, however, that eqn. (5) should be valid for paths that are sufficiently oblique, since the proportion of deviative absorption appears to become less important as the path length is increased. This suggests that better accuracy would result if oblique paths were to be used as standards of comparison rather than vertical-incidence ones as at present. This has not been done in the past because of greater instrumental complication, but it appears to be desirable to do so if greater accuracy in forecasting absorption and lowest usable frequencies is required.

Martyn's theorem can be stated as follows: if the effect of the earth's magnetic field and curvature of the earth can be neglected, the absorption in decibels at frequency f at oblique incidence equals cos i times the absorption in decibels at frequency f cos i at vertical incidence, where i is the angle of incidence for the oblique path. The frequencies f and f cos i are said to be equivalent.

From the present measurements made at a frequency of 1·62 Mc/s at vertical incidence, the absorption at noon to be expected from Martyn's theorem for the oblique equivalent frequency of 5·1 Mc/s is about 10 dB, which is clearly greatly in error, since the measured result was almost certainly above 25 dB. A better result, 37 dB, is obtained if the theorem is modified by omitting the cos i multiplying factor, as suggested by Beynon.[1,3] However, the explanation advanced by him for F-reflections (losses due to partial reflections) is hardly tenable in this instance since the top of the trajectory is below the level at which scattering due to E_s normally occurs. However, it has been shown elsewhere[6] that the presence of the earth's magnetic field increases the absorption calculated for oblique incidence, and it therefore seems unwise to use the theorem in either form except as a very rough guide.

(3.3) Extension of the Results to Other Paths and Periods of Time

From the results of Section 3.1 it is probable that the formula

$$A_i = A \sec i \cos^{1/2}\chi/(f + f_L)^2 \text{ decibels} \quad . \quad . \quad (6)$$

will give the absorption suffered by an ordinary ray of frequency f megacycles per second at angle of incidence i ($i > 70°$) at solar zenith angle χ. The quantity A has been found to be 310 dB × (Mc/s)^2 for summer at sunspot minimum, but the difficulty remains of predicting it for other seasons and periods in the sunspot cycle. In particular, absorption, at least when measured at vertical incidence, appears to be anomalous during winter (November–March) in medium and high latitudes; during

is period it varies greatly in level from week to week and, mpared with eqn. (6), is too high on the average by an amount hich is nearly independent of position in the sunspot cycle. n examination of published and unpublished vertical-incidence ata for 2 Mc/s (at which frequency absorption calculated for blique incidence has been found to agree fairly closely with the easurements incorporated in the present paper) shows that A eqn. (6) may be regarded as being increased during the period ovember–March by this effect, for Slough, according to the npirical formula

$$\Delta A_w = (350 \pm 90) \sin [11 \cdot 2 \text{ (no. of weeks, } \not> 16, \text{ beyond 1st November)}]^\circ \quad (7)$$

)% confidence limits are implied in this and succeeding ormulae where limits are given.

As far as can be ascertained, the effect appears to be due to an hancement of the non-deviative absorption in the lower - or D-regions; consequently the vertical-incidence data from hich this and subsequent formulae are calculated should be pplicable to oblique incidence.

Very approximately, assuming that ΔA_w varies linearly with titude, eqn. (7) becomes

$$A_w = (6 \cdot 8 \pm 1 \cdot 7)|(\text{latitude }^\circ)| \times \sin [11 \cdot 2 \text{ (no. of weeks beyond 1st November)}]^\circ \quad (8)$$

Variation of the normal (non-winter) absorption over the inspot cycle[8] may be expressed, for a frequency of 2 Mc/s at ertical incidence, by the empirical formula

$$A_{sc} = (160 \pm 20) \times \sin^2 [16 \cdot 4 \text{ (no. of years beyond sunspot minimum)}]^\circ \quad (9)$$

The overall variation between sunspot minimum and sunspot aximum during summer and winter can be seen in Table 2 elow.

Table 2

ARIATION OF NOON ABSORPTION AT VERTICAL INCIDENCE AT 2·0 Mc/s

	Absorption at sunspot minimum dB				Absorption at sunspot maximum dB			
	1942	1943	1944		1946	1947	1948	
May		31·6	30·0	33 ± 3·5		49·6	49·5	46 ± 4
June		29·7	34·9			49·0	41·5	
July		33·5	36·5			50·0	45·2	
Dec.	40·0	39·8		33 ± 7	35·2	36·8		39 ± 5
Jan.	34·6	29·4			40·0	43·8		
Feb.	28·4	26·4			40·0	34·5		

The winter effect can be seen particularly well for sunspot inimum, when the mean winter absorption is practically the ame as the summer value. During the summer, most of the ariation between the absorption figures given in Table 2 can be tributed to sampling errors brought about by fading. During e winter, however, increase of absorption due to the effect rmulated in eqn. (7) occurs, which causes some of the variation corded and elevates the mean and median values.

(4) FADING

The procedure described in Sections 2 and 3 will enable the r.m.s. value of the fluctuating signal to be calculated, which proximates to the mean or median value.[4] Consequently a level which is exceeded for only about 50% of the time will be obtained. This percentage is often inadequate for practical purposes, and additional transmitter power is needed to ensure that the required level is exceeded for a higher percentage of the time. A method of enabling this additional power to be estimated, based on the theoretical random statistics[4] for a single-ray transmission, will now be indicated. The standard deviation σ appropriate to rapid fading for vertical incidence ($i = 0°$) has been found[4] to be about 47% of the mean signal level, and a Rayleigh distribution is closely approached. The value for the Slough–Inverness path ($i = 74°$) at 5·1 Mc/s was 28%, and the effects are not likely to vary greatly with frequency. Insufficient data are available to calculate σ as a function of angle of incidence for other paths, but it could be estimated for a particular path from the values just given. For the percentage value of σ so obtained, the signal in decibels (relative to the r.m.s. value of the fading signal) which is actually exceeded for the desired percentage of the time can then be read directly from Fig. 1 of Reference 4. This number of decibels must be made up by additional transmitter power or aerial gain if the original r.m.s. value is to be exceeded for the same time. For example, an extra 18 dB is required at vertical incidence to ensure that the rapid-fading effects are overcome for 99% of the time.

In addition, variations due to slow fading[4] must be considered, a further 8 dB of transmitter power or aerial gain being required to ensure that the calculated level is exceeded for 99% of the time. A distinction between rapid and slow fading is of practical importance, since the variation due to rapid fading can be greatly reduced by conventional diversity methods, whereas it is doubtful whether the effects of slow fading can be. It is probable that slow fading is due to ionospheric tilts, and Bramley[9] has found that the spacing between points on the tilting surface needs to be 50–100 km for the variation of tilt with time at these points to be uncorrelated, at least for F-reflections. If, by implication, the amplitude variations are also uncorrelated at such spacings, the effect of slow fades can only be reduced by diversity reception from points many kilometres apart.

(5) CONCLUSIONS

The results of some measurements of absorption and fading of ordinary rays, using pulses reflected once from the E-region, have been discussed and compared. The measurements were made at vertical incidence at Slough and at oblique incidence over the path between Slough and Inverness.

Absorption calculated for oblique incidence by the 'non-deviative' formula from measurements made simultaneously at vertical incidence at frequencies between 1·6 and 4·8 Mc/s was on the whole too high, although the values calculated from the measurements made at 1·62 and 2·0 Mc/s agreed to within experimental accuracy. Martyn's absorption theorem, on the other hand, gave results much too low, the accuracy being improved by omitting the cos i multiplying factor. These results appear to be in general agreement with the indications of Booker–Millington magneto-ionic theory.

An attempt to generalize the application of the results to other E-region trajectories in, say, the frequency range 1·5–15 Mc/s, at different times of the year and sunspot cycle has been made by incorporating conclusions reached from routine absorption measurements which have been made at Slough at vertical incidence for many years. However, in the absence of further data the generalizations so made are tentative, and it is emphasized that, if work of greater accuracy is required in the future, absorption should, in general, be calculated from measurements made on oblique paths, rather than from those made at vertical

incidence. Simultaneous vertical and oblique measurements might well be used, however, in obtaining results useful in ionospheric physics, and in separating deviative from non-deviative absorption.

(6) ACKNOWLEDGMENT

The work was carried out as part of the programme of the Radio Research Board. The paper is published by permission of the Director of Radio Research of the Department of Scientific and Industrial Research.

(7) REFERENCES

(1) BEYNON, W. J. G.: 'Some Notes on the Absorption of Radio Waves reflected from the Ionosphere at Oblique Incidence', *Proceedings I.E.E.*, Paper No. 1569 R, January, 1954 (**101**, Part III, p. 15).

(2) ALLCOCK, G. MCK.: 'Ionospheric Absorption at Vertical and Oblique Incidence', *ibid.*, Paper No. 1664 R, April, 1954 (**101**, Part III, p. 360).

(3) APPLETON, E. V., and BEYNON, W. J. G.: 'An Ionospheric Attenuation Equivalence Theorem', *Journal of Atmospheric and Terrestrial Physics*, 1955, **6**, p. 141.

(4) MEADOWS, R. W., and MOORAT, A. J. G.: 'The Effect of Fading on the Accuracy of Measurement of Ionospheric Absorption' (see page 27).

(5) ECKERSLEY, T. L., and MILLINGTON, G.: 'The Limiting Polarization of Medium Waves reflected from the Ionosphere', *Proceedings of the Physical Society*, 1939, **51**, p. 110.

(6) MEADOWS, R. W., and MOORAT, A. J. G.: 'The Effect of the Earth's Magnetic Field on Absorption for a Single-Hop Ionospheric Path' (see page 33).

(7) PIGGOTT, W. R.: Discussion on Ionospheric Absorption and Reflection of Radio Waves, *Proceedings I.E.E.*, 1954, **101**, Part III, p. 370.

(8) 'Monthly Bulletins of Ionospheric Characteristics', Department of Scientific and Industrial Research.

(9) BRAMLEY, E. N.: 'Some Comparative Directional Measurements on Short Radio Waves over Different Transmission Paths', *Proceedings I.E.E.*, Paper No. 1716 R, September, 1954 (**102 B**, p. 544).

1.396.11

The Institution of Electrical Engineers
Paper No. 2428 R
Jan. 1958

THE EFFECT OF FADING ON THE ACCURACY OF MEASUREMENT OF IONOSPHERIC ABSORPTION

By R. W. MEADOWS, B.Sc.(Eng.), Associate Member, and A. J. G. MOORAT.

(*The paper was first received* 1st *January, and in revised form* 12th *July*, 1957.)

SUMMARY

The rapid and slow components of a fading signal are separated by -emi-empirical process in which each component is, in turn, assumed .consist of a steady or specularly reflected component with a random mponent added. The process is applied to some measurements of : amplitude of first-order reflections from the E-region at vertical :l oblique incidence on equivalent frequencies. The standard viation of the amplitude variation due to rapid fading was found to greater at vertical than at oblique incidence, but insufficient evidence yet available to determine whether the variation due to slow fading s also greater.

The accuracy of measurement of the smoothed value of field strength defined as the range of values having a 99% chance of containing : correct value. On this basis, the accuracy of noon absorption, as culated from a single day's observations at one frequency, has been imated to be about +4, −12 dB at vertical incidence and +4, −11 dB oblique incidence. These limits are not symmetrically disposed out the mean value, as is conventional in normal statistics, partly cause of the effect of the decibel scale, and partly owing to the aracteristics of deep fading.

(1) INTRODUCTION

It is well known that the field strength of pulses reflected from particular level in the ionosphere is subject to great amplitude riations, known as fading, which appear to be of two kinds, oid and slow. These variations complicate the measurement the average power loss due to electron collisions (absorption) ffered by waves travelling through the ionosphere, since this volves estimating the steady value of the field strength which ould have been obtained had no fading occurred. The absorp-n at a particular time of day is taken as the ratio of the oothed value of field strength occurring under conditions of ss-free transmission to that occurring at the time in question. is ratio is greater than unity and is usually expressed in cibels.

The type of smoothing used is discussed in the paper and the curacy with which the smoothed value approaches the steady ue of field strength is assessed. In order to do this, a method separating rapid and slow fading is evolved.

Although it is of importance in estimating how closely a rticular theory fits the measured results, little work on this oject appears to have been published, probably because of its angibility. The authors have themselves found difficulty in mulating the subject, and are conscious of some limitations the treatment which render it semi-empirical.

(2) DETERMINATION OF STEADY AMPLITUDE FROM RAPID-FADING STATISTICS

The period of rapid fading is usually between about 0·5 and sec, whereas that of slow fading is between 10 min and one ur. or more. Both are here assumed to be random in racter, but to have different time-scales. Rapid fading is considered to be due to irregular patches of ionization drifting across the ray path,[1] whilst the slow fading is due to alteration of curvature and tilt of the reflecting surface as a whole.[2,3]

The discussion which follows in Sections 3 and 4 will deal primarily with the rapid fading. Slow fading will then be regarded, in Section 5, as causing the perturbations which remain when all the variation which might reasonably be expected from the rapid fading and change in absorption is removed.

McNicol[4] considers that a rapidly fading echo may be regarded as consisting of a steady or specularly reflected component with randomly scattered components added, the resultant amplitude being statistically calculable by a formula, due to Rice,[5] derived originally for a sinusoidal voltage with added noise. Now, for a fading echo it seems axiomatic that, as the proportion of energy returned by such scattering ($\overline{x^2}$, say) increases, so the energy returned specularly (s^2, say) must decrease. If it is further assumed that $s^2 + \overline{x^2}$ remains constant whatever fading depth occurs, then the r.m.s. value of a succession of echo-amplitude measurements should give the value of the signal which would be obtained if the ionosphere were steady, i.e. with $x = 0$; the r.m.s. value might then be expected to approximate better to the true undisturbed value than does the arithmetic mean.

The deepest fading obtainable would occur when the specular value was zero, since only the fluctuations then remain. Then if $\bar{R}$ is the arithmetic mean value of the received echo, averaged over a period of time, and $\overline{R^2}$ is the corresponding mean-square value, it can be shown that $\overline{R^2}/\bar{R}^2 = 4/\pi$, which represents the maximum possible value of the ratio provided that the fading is truly random. The distribution of amplitudes is then of the Rayleigh type.[1] Also, if σ is the standard deviation of the succession of amplitude samples of the fading signal, then generally, $\overline{R^2} = \bar{R}^2 + \sigma^2$, the r.m.s. value always being greater than the mean. From these two equations, $\sigma = 52 \cdot 5\%$ of the mean, or 50% of the r.m.s. value of the individual amplitude measurements. This evidently represents the maximum value it can attain; lower values are considered later when Rice's theory is extended. However, the maximum error in taking the mean value in place of the r.m.s. value is $\sqrt{\pi}/2$ in magnitude, or about 1 dB, and is therefore small.

(3) A DEFINITION OF ACCURACY OF MEASUREMENT OF A FADING ECHO

When stating the measured value of a quantity it is usual to specify the accuracy of that measurement by assigning limits between which the true value stands a definite chance of lying. For present purposes this chance will be taken as 99 : 1, which implies that 99% of the instantaneous measurements of the amplitude of a particular fading echo taken over a particular period of time might then be expected, on the average, to lie within the limits specified.

For small percentages of random component, σ, superimposed upon a fixed specular value, s, the mean value of the resultant of the combination remains very close to s, since the moment of the variation occurring above s cancels the moment below it;

Wri:en contributions on papers published without being read at meetings are ted for consideration with a view to publication.

The paper is an official communication from the Radio Research Station, Department f Scientific and Industrial Research.

the distribution is, in fact, substantially Gaussian. Under these conditions 99% of the variation would be contained within limits placed symmetrically about s, and the accuracy of representation of the resultant could be specified in the form $s + r$. The r.m.s. value of the resultant would also lie very close to the mean value. However, for the large variations that commonly occur with fading signals, this symmetry of variation no longer holds, since the distribution of amplitudes becomes skew. Skewness occurs because amplitudes can rise indefinitely but cannot fall below zero; the downward-going variations are therefore compressed. However, when expressed logarithmically, e.g. in decibels below the mean or r.m.s. value, the apparent downward variations appear to be expanded. Ranges of values containing 99% of the observations can again be specified, but still will not lie symmetrically around the mean value. The 99% range adopted in the present paper lies between a value exceeded for 99·5% of the time and one exceeded for 0·5% of the time.

It is convenient to express the variability of the fading signal in terms of the ratio of the r.m.s. to mean values of a succession of amplitude samples, both of these quantities being readily computed from the observations. A family of curves, Fig. 1, has

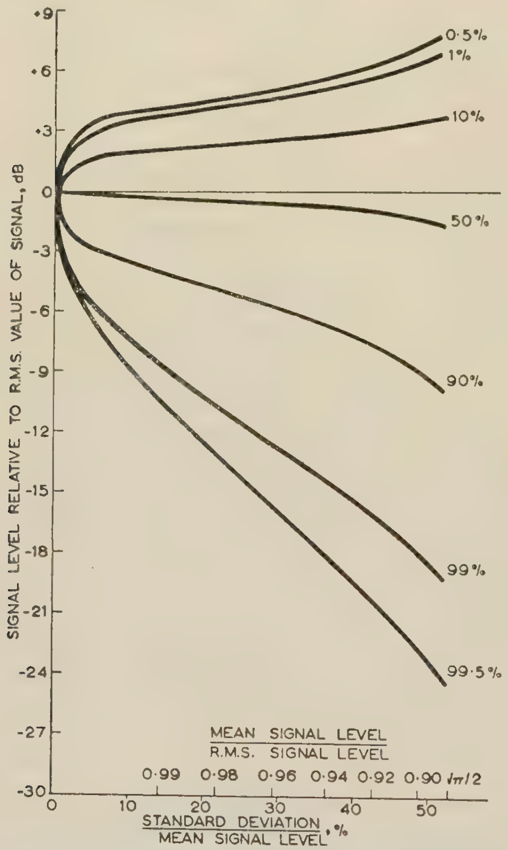


Fig. 1.—Theoretical characteristics for random fading.

Signal level exceeded for percentage of time indicated.

been prepared showing the signal level, in decibels, exceeded for various percentages of the time as a function of the variability, assuming that the amplitudes obey a Rice distribution. In particular, the 99·5% and 0·5% curves are of interest for determining the 99% range. The corresponding ratios between the specular and the fluctuating components can be obtained using Fig. 2. The method of calculating these curves is indicated in Section 9.

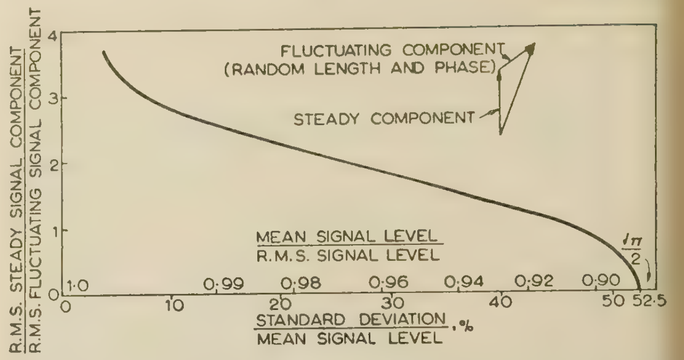


Fig. 2.—Components of fading signal in terms of observable quantities.

(4) APPLICATION TO SOME VERTICAL-INCIDENCE MEASUREMENTS

Special measurements of the (relative) amplitude of first-order E-echoes were made at 1·62 Mc/s by a standard method.[6,7] These measurements extended over several 24-hour periods from 16th–30th June, 1954; the night-time values were obtained in order to calibrate the equipment in absolute terms by using ratios of the strength of multiple reflections.[6,7] Usually a measurement of the amplitude of the fading echo was made about once every 10 sec, since it is known that readings taken at this separation are substantially uncorrelated samples of the short-term distribution.

The observations were divided into blocks of 10. The mean value of each block and the standard deviation of the 10 observations from it were evaluated. To reduce the effects of slow variations, each of these values of standard deviation was expressed as a percentage of the mean value about which it was calculated.

The values so obtained are plotted in Fig. 3 for one typical day, and no substantial variation of the mean level of the results is apparent throughout the day, although the variation from minute to minute is great. The mean standard deviation for the whole day is 47%, which is approaching the maximum theoretical figure of 52·5% mentioned in Section 2.

Since the successive amplitude readings are considered to be uncorrelated on a short-term basis, the standard deviation of the mean ('standard error' or s.e.) of an individual block of 10 is $47/\sqrt{10} = 15\%$ of the mean amplitude. This value may be compared with that obtained by Beynon and Davies[8] which, from their Fig. 1, is about 34%. The fact that their value is greater is significant. It is based on the mean value of 200 observations taken over a period of one hour just after noon, when true absorption changes are normally least. This period is, however, certainly long enough for some slow fading to have taken place, so that their statistics refer to the combined effects of slow and rapid fading. It is also probable that some change in absorption occurred, since their measurements were made at the equinoxes when the rate of change of absorption with time around noon is greatest. This is confirmed by their Fig. 2.

From Fig. 1 it will be seen that for a standard deviation of 15% of the mean there is a 99% chance of finding the r.m.s. value of the whole distribution between approximately 11 dB below the r.m.s. value and 4·5 dB above it, i.e. within a range of 15·5 dB. This is therefore the accuracy of measurement, as defined in Section 3, for the mean of samples taken over a period of about one minute.

The mean values of the blocks of 10 successive observations as recorded (i.e. uncorrected for slow fading by the process described above) were plotted against time on a convenient decibel scale. A pair of lines spaced by the above-mentioned

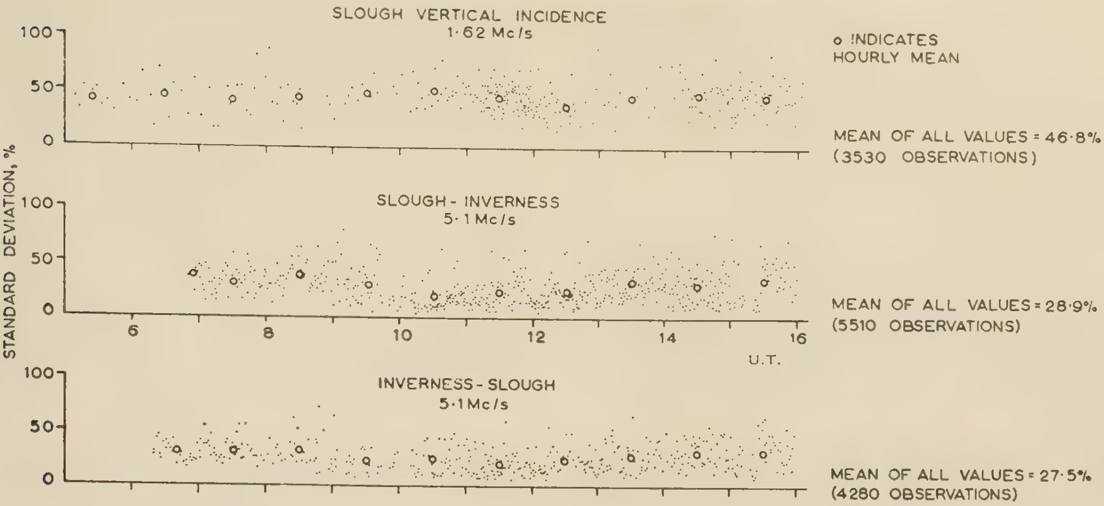


Fig. 3.—E-region fading-depth statistics.

Standard deviations of blocks of 10 successive instantaneous observations of amplitude (expressed as percentage of the mean value of each block, to eliminate slow variations), 30th June, 1954, covering solar eclipse.

·5 dB was then drawn to enclose the majority of the mean ues throughout the day. This pair of lines is shown dashed Fig. 4; the individual mean values are too numerous to be tted satisfactorily and have therefore been omitted. A curve wn 11 dB above the lower dashed line (or 4·5 dB below the per one) then represents the field strength smoothed to elimi- te the effects of rapid fading. This has been shown as a ain-dotted line in Fig. 4. Variations due to slow fading and rnal change of absorption remain. The diurnal change of sorption will be assumed to be symmetrical about noon—an umption which is consistent with the experimental results. ch a symmetrical curve giving the best fit of the chain-dotted e has been drawn as a full line in Fig. 4, and therefore represents most probable manner in which the absorption changes oughout the day. Differences between the full and chain- tted lines must be regarded as being due to slow fading on the esent argument; in particular, a pronounced slow variation er noon is apparent.

One day's observations is, however, insufficient for a statistical alysis of the slow fading to be made, since there may be preciable correlation between successive values as far apart as e hour. In order, therefore, to form an idea of the variations e to slow fading which are likely to occur, the values of sorption measured (as a matter of routine) each day at Slough noon at a frequency of 2 Mc/s were examined.

It was found that the mean values at noon for May, June and y, 1954, were, respectively, 29·7, 28·3 and 30·1 dB.* Estima- n of the standard deviation of these noon values taken over the nonth period showed that the difference between the monthly an values was hardly significant statistically and was attri- table to the spread of the random distribution of the individual on values. The implication is that the variation of absorption er the period due to alteration of the solar zenith angle was gligible, a result which is to be expected during the summer sunspot minimum when the observations were made. This confirmed by comparisons of values of absorption measured different frequencies throughout a sunspot cycle,[9] when it s found that the majority of the day-to-day changes occurring

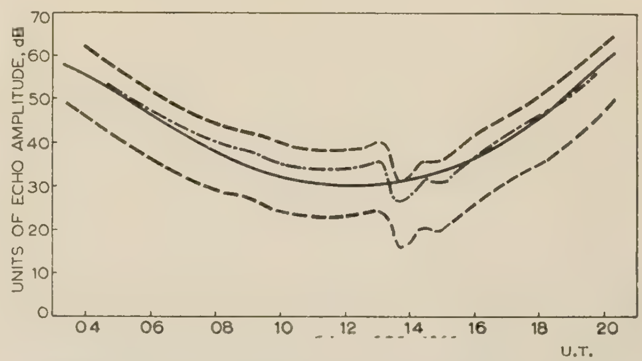


Fig. 4.—Field strength measured at vertical incidence.

——— 15½ dB spread of 1 min means due to rapid fading.
— · — Variation remaining after rapid fading is removed.
——— Estimated variation due to change in absorption.

at a single frequency were due to fading phenomena. Consequently, it will be assumed that the noon-to-noon variability was entirely due to a combination of rapid and slow fading.

The extreme values of the distribution were 40·4 and 22·2 dB, with about 80% of the values lying between limits of about 5 dB below and 2·5 dB above the r.m.s. value. Fig. 1 shows this to correspond to a standard deviation of 19% of the mean. This figure is based on a rather small number of noon observations, but it should, nevertheless, give a rough idea of the variability due to slow fading.

Each noon value is the mean of some 50 individual measurements spread over about 12 min, a period which is reasonably short compared with the slow fading at this frequency. Thus the standard error due to rapid fading is $0{\cdot}47/\sqrt{50} = 0{\cdot}07$. The standard deviation of the slow fading causing the noon-to-noon variations is therefore $\sqrt{(0{\cdot}19^2 - 0{\cdot}07^2)} = 17\%$, corresponding to a range, containing 99% of the observations, lying between 11 dB below and 4 dB above the r.m.s. value (Fig. 1). This figure would correspond to an 'accuracy of measurement' of the r.m.s. value of +4, −11 dB, and represent the limiting accuracy imposed by slow fading on the r.m.s. value likely to be obtained if no fading of any kind were present.

It is evident that the variations due to slow fading are so large that the accuracy of assessing the value of absorption by

* It may be noted here that the arithmetic mean of a succession of decibels leads the geometric mean of the linear values which the decibels express logarith- all , and the geometric mean is always lower than the arithmetic mean. However, r maximum variability corresponding to a Rayleigh distribution the discrepancy to this cause is small. Except when otherwise stated, s.d., s.e., mean and r.m.s. es refer to the original linear observations and not to the decibel values expressing el of those observations.

taking the mean of a number of rapid observations increases very slowly with the number of measurements. This is demonstrated in Table 1, where the resultant limitation of accuracy by combined rapid and slow fading is given. Two separate standard deviations for slow changes are assumed, namely 17% and 7%.

The limiting accuracy has been approached after 10 min of measuring in the presence of the larger slow fade and after 20 min with the smaller.

The period around noon available for assessing the noon field strength, during which time the absorption should not have changed appreciably due to alteration of the solar zenith angle, is about two hours at the most in summer and winter, but less during the equinoxes. Owing to the extended time-scale of slow fading, it is doubtful whether more than two combined samples (Table 1) could with certainty be regarded as uncorrelated; in consequence, the limits which define a 99% probability of containing the r.m.s. value of the (undisturbed) field would correspond to a standard error of $0{\cdot}17/\sqrt{2}$, i.e. (from Fig. 1) +4 and −10 dB. This is therefore the probable accuracy of measurement of the r.m.s. value of the noon field strength for a single frequency.

The accuracy of assessing the unabsorbed field from night-time measurements is somewhat greater, since at night the absorption is low and multiple reflections can usually be recorded. During the night-time periods of the sequence of measurements being discussed, recordings of 1F, 2F and 5F echoes were made. Since the absorption during the night should remain constant for several hours, it is probable that about five uncorrelated long-term values of each were permissible, say 15 altogether. The standard error is therefore $0{\cdot}17/\sqrt{15} = 4{\cdot}4\%$, corresponding to an accuracy of measurement of the r.m.s. value of field strength of +3, −6 dB.

The absorption at noon is defined as the ratio of noon field to unabsorbed field; the accuracy of the ratio will be evaluated from a standard error equal to the square root of the sum of the squares of the standard errors of the two quantities comprising it. The value of the ratio was found to be 34 dB, with 99% confidence limits of 30 dB and 46 dB. Instrumental errors were not considered significant because, for relative measurements of field strength taken over periods which are not too long, they are usually much smaller than sampling errors.

(5) APPLICATION TO SOME OBLIQUE-INCIDENCE MEASUREMENTS

Measurements of the amplitude of one-hop E-region echoes received at Slough from a pulsed transmitter at Inverness were made at approximately 10 sec intervals during the period 23rd June to 7th July, 1954. During about 75% of the time, similar measurements on E-region echoes received at Inverness from Slough were also made, using the same aerials. The frequency was maintained as close to 5·1 Mc/s as possible, but occasionally had to be altered by a few kilocycles per second in one direction or the other to avoid interference.

The rapid fading was again assessed using the process of evaluating the percentage standard deviation of amplitude-samples taken at 10 sec intervals (Section 4). The results are shown in Fig. 3. The mean standard deviation for the whole day was 27·5% for the Inverness to Slough direction and 28·9% for Slough to Inverness, the difference being attributable entirely to sampling errors. The standard deviation of the mean of 10 uncorrelated observations comprising a block is now 9%, compared with 15% obtained by the same process for vertical incidence in Section 4. From Fig. 1, a standard deviation of 9% of the mean implies that 99% of the observations lie within +4 dB and −8 dB of the r.m.s. value. These figures would correspond to an accuracy of measurement of the r.m.s. value of +4, −8 dB, a total range of 12 dB, if rapid fading only were present.

A pair of lines indicating this 12 dB spread has been drawn in Fig. 5 for a typical day, thus showing up the significant

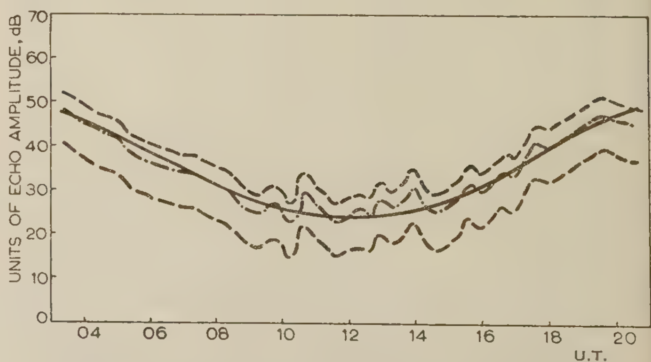


Fig. 5.—Field strength measured at oblique incidence.

——— 12 dB spread of 1 min means due to rapid fading.
— · — Variation remaining after rapid fading is removed.
——— Estimated variation due to change in absorption.

variations due to slow fading in a way similar to Fig. 4 for the vertical-incidence case. The number of noon observations is again not sufficiently large to enable the statistics of these long-term variations to be assessed. However, comparing the oblique with the vertical-incidence results, it seems that the smaller

Table 1

THE AVERAGE EFFECT OF COMBINED RAPID AND SLOW FADING ON IMPROVING THE ACCURACY OF MEASURING R.M.S. FIELD STRENGTH AT VERTICAL INCIDENCE BY INCREASING THE NUMBER OF OBSERVATIONS

Number of observations at 10 sec intervals n	Time taken	Standard error due to rapid fading $0.47/\sqrt{n}$	Standard deviation due to slow fading	Resultant standard deviation	Accuracy of r.m.s. value (99% confidence)
	min				dB
1	0·2	0·47	0·17	0·50	+7, −23
			0·07	0·48	+ 7, −22
4	0·7	0·23	0·17	0·29	+5, −16
			0·07	0·24	+ 5, −14
16	2·6	0·12	0·17	0·21	+5, −13
			0·07	0·14	+ 4, −10
64	10·7	0·06	0·17	0·18	+4, −11
			0·07	0·09	+ 4, − 9
128	21·4	0·04	0·17	0·17	+4, −11
			0·07	0·08	+ 4, − 8

ort-term variations at oblique incidence disclose more of the tail of the long-term variations which, in consequence, cannot demonstrated to be less than those at vertical incidence. It il therefore be assumed that long-term variations are at least great, i.e. unlikely to vary by more than +4 and −11 dB om the diurnal absorption line for more than 1% of the time. ie average accuracy of measurements due to combined rapid square root of the sum of the squares of the standard errors of the two values of field strength. The accuracy of measurement of the absorption value at noon then becomes about +4, −11 dB. The most probable value itself was found to be 29 dB. Thus, using the observations from a single day, it has been found that the absorption at oblique incidence has a 99% chance of lying between 25 and 40 dB.

Table 2

IE AVERAGE EFFECT OF COMBINED RAPID AND SLOW FADING ON IMPROVING THE ACCURACY OF MEASURING ABSORPTION AT OBLIQUE INCIDENCE BY INCREASING THE NUMBER OF OBSERVATIONS

Number of observations at 10 sec intervals n	Time taken	Standard error due to rapid fading $0.28/\sqrt{n}$	Standard deviation due to slow fading	Resultant standard deviation	Accuracy of r.m.s. value (99% confidence)
	min				dB
1	0·2	0·28	0·17	0·33	+5, −17
			0·07	0·30	+ 5, −16
4	0·7	0·14	0·17	0·22	+5, −13
			0·07	0·16	+ 4, −11
16	2·6	0·07	0·17	0·18	+4, −12
			0·07	0·10	+4, − 9
64	10·7	0·03	0·17	0·17	+4, −11
			0·07	0·08	+ 4, − 8
128	21·4	0·02	0·17	0·17	+4, −11
			0·07	0·07	+ 3, − 7

d slow fading is given in Table 2. As at vertical incidence, o standard deviations for the slow fading are used, namely % and 7% of the mean.

The limiting accuracy has been approached after 10 min of asuring in the presence of the larger slow fade and after min with the smaller. The approach to this accuracy is so w that, except for special purposes, there will be little point averaging over periods longer than about 5 min.

t is doubtful whether more than about two long-term samples packets of rapid observations) can on the average be regarded uncorrelated around noon, and the probable accuracy of imating the value of the field strength around noon would be 4, −10 dB.

During the night the strength of reflections from E_s oradic-E) occurring at about the same height as reflections m normal-E during the day, were recorded. This was neces-y since normal E-reflection does not in general, and in fact did , persist throughout the night. Multiple E_s reflections were ely observed, and reflections from the F-region, occurring about 250 km during most of the night, were not used ce the wave angle did not correspond to the day-time E-value. s not, therefore, possible to estimate the unabsorbed field as vertical incidence, since ratios of the strength of multiple ections cannot be used. Consequently, the night-time values ained from sporadic-E reflections will be assumed to equal unabsorbed values. This assumption leads to a value of on absorption which agrees with absolute measurements,[10] d, in addition, the measurements made simultaneously at 2 Mc/s on vertical incidence were found to approach the absorbed value closely.

t was considered possible that five long-term values could be arded as uncorrelated throughout the night, assuming that ections from sporadic-E were subject to the effects of slow ing as well as rapid in the same general manner as normal-E. ere does not appear to be any evidence for this, but it is not an easonable supposition and it results in a standard error of %, leading to 99% confidence limits of +4, −8 dB. The orption at noon is the ratio of the noon field to the unabsorbed d calculated from night-time values; as before, the accuracy the ratio will be evaluated from a standard error equal to the

(6) CONCLUSIONS

A semi-empirical statistical method of assessing the accuracy of ionospheric absorption calculated from amplitude measurements of a fading echo has been outlined. The rapid fading has been separated from the slow fading and the statistics of each have been evaluated. The standard deviation of the amplitude variation due to rapid fading was found to be greater at vertical incidence than at oblique, but insufficient evidence is yet available to determine whether the variation due to the slow fading was also greater.

The accuracy of measurement of the r.m.s. value has been defined as the range of values having a 99% chance of including the correct value; the range therefore represents 99% confidence limits. On this basis, the accuracy of noon absorption as calculated from a single day's observations at one frequency has been estimated to be about +4, −12 dB at vertical incidence and +4, −11 dB at oblique. These limits are not symmetrically disposed about the mean value, as is conventional in normal statistics, partly because of the effect of the decibel scale and partly owing to the characteristics of deep fading.

(7) ACKNOWLEDGMENT

The work was carried out as part of the programme of the Radio Research Board. The paper is published by permission of the Director of Radio Research of the Department of Scientific and Industrial Research.

(8) REFERENCES

(1) RATCLIFFE, J. A.: 'Diffraction from the Ionosphere and the Fading of Radio Waves', *Nature*, 1948, **162**, p. 9.

(2) MUNRO, G. H.: 'Short-Period Changes in the F-Region of the Ionosphere', *ibid.*, 1948, **162**, p. 886.

(3) MARTYN, D. F.: 'Cellular Atmospheric Waves in the Ionosphere and Troposphere', *Proceedings of the Royal Society*, A, 1950, **201**, p. 216.

(4) MCNICOL, R. W. E.: 'The Fading of Radio Waves of Medium and High Frequencies', *Proceedings I.E.E.*, Paper No. 891 R, November, 1949 (**96**, Part III, p. 517).

(5) RICE, S. O.: 'Statistical Properties of Random Noise', *Bell System Technical Journal*, 1945, **24**, p. 46.

(6) PIGGOTT, W. R.: Discussion on Ionospheric Absorption and Reflection of Radio Waves, *Proceedings I.E.E.*, 1954, **101**, p. 370.

(7) PIGGOTT, W. R.: 'D.S.I.R. Ionospheric Absorption Measuring Equipment', *Wireless Engineer*, 1955, **32**, p. 164.

(8) BEYNON, W. J. G., and DAVIES, K.: 'Simultaneous Ionospheric Absorption Measurements at Widely Separated Stations', *Journal of Atmospheric and Terrestrial Physics*, 1954, **5**, p. 273.

(9) APPLETON, E. V., and PIGGOTT, W. R.: 'Ionospheric Measurements through a Sunspot Cycle', *ibid.*, 1954, **5**, p. 141.

(10) MEADOWS, R. W.: 'The Attenuation of Radio Waves reflected from the E-Region of the Ionosphere' (see page 22).

(9) APPENDIX

Determination of Level exceeded for a given Percentage of the Time for Fading Echoes obeying a Rice Distribution of Amplitudes

The statistics of the Rice[5] distribution may be expressed in terms of the mean and mean-square values of a succession of amplitude measurements comprising it, so that, in Rice's notation,

$$\overline{R^n} = (2\psi_0)^{n/2}\Gamma\left(\frac{n}{2}+1\right)\varepsilon^{-P^2/2\psi_0}{}_1F_1\left(\frac{n}{2}+1; 1; \frac{P^2}{2\psi_0}\right)$$

where $\overline{R^n}$ is the mean value of the sum of the nth powers of the succession of values (samples) of the instantaneous amplitude of the fading signal envelope R, comprising a sinusoidal component of (peak) amplitude P and random fluctuations of r.m.s. value $\sqrt{\psi_0}$. The gamma function Γ and hypergeometric function ${}_1F_1$ can be found in reference tables.*

* JAHNKE, E., and EMDE, F.: 'Tables of Functions' (B. G. Teubner, Leipzig, 1938), p. 275.

The ratio of the mean ($n = 1$) and r.m.s. ($n = 2$) values then becomes

$$\frac{\bar{R}}{\sqrt{\overline{R^2}}} = \frac{\text{mean}}{\text{r.m.s.}} = \frac{\sqrt{\pi}}{2\sqrt{(1+b^2)}}{}_1F_1\left(\frac{3}{2}; 1; b^2\right)$$

where $b = P/\sqrt{(2\psi_0)} = \dfrac{\text{r.m.s. specular component}}{\text{r.m.s. fluctuations}}$.

A plot of b against the ratio mean/r.m.s. is given in Fig. 2, where the values of standard deviation, σ, corresponding to ratio of mean to r.m.s. values are scaled, according to the formula

$$\left(\frac{\sigma}{\bar{R}}\right)^2 = \frac{\sqrt{R^2}}{\bar{R}} - 1$$

Let $a = b\sqrt{2}$. Then a value of a corresponding to a particular value of mean/r.m.s. ratio for a particular fading distribution is used to specify a curve in Rice's Fig. 7 (Reference 5, p. 103) giving the percentage of time for which the fading signal is less than the r.m.s. fluctuation level; these percentages have then, of course, to be subtracted from 100 to give the percentage of time for which the signal exceeds the r.m.s. fluctuation level.

The value of $v - a$ (Rice's notation, where $v = R/\sqrt{\psi_0}$, corresponding to the ordinate expressing the desired percentage of the time the fading signal exceeds the r.m.s. fluctuation level is then read off the curve; hence v is determined and the level by which the resultant exceeds the r.m.s. value of the distribution (as opposed to the r.m.s. fluctuations) for the same percentage of the time is obtained from

$$\frac{R}{\sqrt{\overline{R^2}}} = \frac{v}{\sqrt{[2(1+b^2)]}}$$

This formula merely refers the r.m.s. level of the envelope to the r.m.s. level of the fluctuations, both quantities being fixed for a given distribution.

Curves showing the level exceeded for various percentages of the time as calculated by the above method are plotted in Fig. 1.

The Institution of Electrical Engineers
Paper No. 2429 R
Jan. 1958
©

THE EFFECT OF THE EARTH'S MAGNETIC FIELD ON ABSORPTION FOR A SINGLE-HOP IONOSPHERIC PATH

By R. W. MEADOWS, B.Sc.(Eng.), Associate Member, and A. J. G. MOORAT.

(*The paper was first received* 1*st January, and in revised form* 12*th July*, 1957.)

SUMMARY

Magneto-ionic calculations show that deviative absorption is not essarily negligible at vertical incidence for waves reflected from the egion at frequencies considerably below the penetration value. nsequently, the value of absorption calculated by the conventional n-deviative' formula for a short-wave oblique path from vertical-idence absorption measurements tends to be too high. Deviative orption on paths sufficiently oblique is, however, negligible. The culations also show the effect of the earth's magnetic field on rtyn's absorption theorem to be similar to the effect of losses due partial reflections: namely to make the absorption in decibels calated from an oblique path from vertical-incidence measurements low by a multiplying factor approaching the cosine of the angle incidence. It is suggested that the absorption to be expected on a io path might best be calculated by applying the conventional n-deviative formula to measurements made at oblique rather than vertical incidence.

(1) INTRODUCTION

Absorption is the term defining 'frictional' losses due to ctron collisions suffered by a ray passing through the iono-here, and its effect is therefore akin to that of resistance in bles and waveguides: the absorption coefficient (in decibels kilometre) is the real part of the propagation coefficient.

Two methods of assessing oblique-path absorption from tical-incidence data have been used in practice for frequencies ween about 3 and 30 Mc/s. The first assumes 'deviative' orption (the absorption introduced during the process of bending in the vicinity of the reflection level) to be small, so t the absorption at frequency f is then proportional to the gth of (undeviated) path through the absorbing D-region, l is usually approximated by[1]

$$\text{Absorption} = \frac{A}{(f \pm f_L)^2} \sec i \ \text{ decibels} \quad . \quad . \quad (1)$$

ere i is the angle of incidence at entry to the ionosphere, l f_L is the gyro-frequency[9] corresponding to the component of gnetic field parallel to the direction of propagation. The itive sign refers to ordinary ray propagation and the negative extraordinary. The paper will be concerned only with ordinary s, since for the frequencies employed they are attenuated less n the extraordinary. A is a constant, numerically equal to absorption in decibels for $f + f_L = 1$, evaluated from ical-incidence ($\sec i = 1$) measurements. The question with method is: to what extent is it really justifiable to neglect deviative absorption present in both the oblique- and vertical-idence cases, especially for engineering purposes when an mate accuracy of about 6 dB is often desirable?

The second method is an application of Martyn's absorption orem, i.e. the absorption in decibels at frequency f at oblique idence is equal to $\cos i$ times that at frequency $f \cos i$ at vertical incidence, *provided that the effect of the earth's magnetic field may be neglected.* The frequencies f and $f \cos i$ are said to be 'equivalent'. The value of absorption at oblique incidence calculated by this method from vertical-incidence values is usually too low, however, and better results are sometimes obtained by omitting the $\cos i$ multiplying factor. Beynon[2] justifies this for trajectories involving F-region reflections by saying that the loss due to partial reflections from sporadic-E ionization on the oblique path, which increases as $\cos i$ decreases, offsets the effect of this term. Allcock's measurements,[5] on the other hand, appear to indicate that a slightly better result is obtained when the $\cos i$ factor is used; but in any case the difference between the mean vertical- and oblique-incidence absorptions on the high-angle path he used turned out to be so small (6 dB in May, 1950) as to have little practical significance. The margin between leaving $\cos i$ in and taking it out was small, perhaps hardly outside the realm of experimental error for absolute measurements.[3]

However, another, or additional, reason for the discrepancy might well be that the earth's magnetic field cannot be neglected. This possibility was strongly supported when measurements[3,4] made by the present authors on oblique rays reflected from the E-region—in which no partial reflection losses should have occurred—agreed better with vertical-incidence measurements when the $\cos i$ term was omitted.

The effect of the earth's magnetic field on the accuracy of the two methods is the subject of the present paper. The work is really the outcome of applying Millington–Booker magneto-ionic theory in an attempt to explain measurements of absorption on a 740 km north-south path between Slough and Inverness, described elsewhere,[3] and already referred to in the paragraph above; but the general implications of the results seemed of sufficient practical importance to warrant reporting separately.

(2) METHOD OF CALCULATING ABSORPTION

Millington[6] has recently developed Booker's oblique-incidence magneto-ionic theory to a form more suited to computation. In Booker's method,[7,8] a ray path, either ordinary or extra-ordinary, is traced through the ionosphere, the characteristics of which may be defined graphically. To simplify the mathematics, however, these characteristics are allowed to vary in the vertical direction only; the ionosphere is then said to be horizontally stratified. For the short oblique path (740 km) being considered, a plane-earth approximation was used.

Millington's method enables the fundamental quartic equation, expressing the vertical component, Z (Millington's notation), of the refractive index, μ, as a function, ζ, of ionization density, to be plotted for a given angle of incidence by introducing an intermediate variable v. The four solutions of this equation represent the upgoing and downcoming ordinary and extra-ordinary rays. Z is first calculated for various values of v, taking the angle of incidence, earth's magnetic-field vector, and wave frequency as parameters. The ionization density function ζ is also calculated for the same range of v as a different function of

itten contributions on papers published without being read at meetings are d or consideration with a view to publication.
e paper is an official communication from the Radio Research Station, rtment of Scientific and Industrial Research.

these parameters. Z is then plotted against ζ for each value of v. So far, no actual ionization-density/height distribution curve has been introduced; the Z/ζ curve relates only to these initial parameters. Measured values of ζ for each physical height z then give Z for each height level considered. The components of ray direction at any height are next evaluated; in general, there will be lateral deviation from the plane of propagation, exept for propagation in the magnetic meridian, which is considered in the present paper. The reflection point, which is always displaced towards the nearer magnetic pole,[8] then remains in the magnetic meridian, but is displaced towards the north in this case. The path of the ray is then determined by graphical integration.

For a point-to-point calculation a guess at the angle of elevation which is necessary to make the ray land on the desired point has to be made; and if the first trial is incorrect, the process has to be repeated until the point is approached to the desired accuracy.

The first approximation to this angle of incidence can, however, be estimated by using[9] Martyn's and Breit and Tuve's theorems. These state that, with no magnetic field, the height of penetration for frequency f at angle of incidence i will be the same as at vertical-incidence frequency $f \cos i$. A graphical construction based on a routine vertical-incidence h'/f record then enables the equivalent frequency, and hence the angle of incidence, to be obtained.

Having determined the angle of incidence suitable to the path, and the ionospheric electron densities, the absorption coefficient, K, at a sufficient number of levels is evaluated, taking the mean collisional frequency appropriate to each level. Graphical integration then gives the total absorption for the particular ray considered, up to the reflection level at which the group path becomes horizontal.

As the reflection level is approached, the rate of change of absorption coefficient with height increases rapidly, becoming infinite at the reflection level itself. It is then difficult to assess the area under the curve, so, following Booker,[7] the neighbourhood of this infinity was approximated by

$$K = \frac{A_0}{(z_0 - z)^n} (n < 1 \simeq 0\cdot5)$$

where A_0 is a constant and z_0 is the height of the reflection level. It was then found helpful to regard the ionization-density/height distribution as a series of straight lines, so that $\zeta = a + bz$, where a and b are constants. Further, over each small range of height being thus considered, in the neighbourhood of the reflection level the collisional frequency f_c may be regarded as being constant. Millington's expression (4) in the second paper of Reference 6 is then

Absorption coefficient $K = \frac{f_c}{2c} f(\zeta)$

So that the total absorption $= \int \frac{f_c}{2c} f(\zeta) dz = \frac{f_c}{2c} \int f(\zeta) \frac{dz}{d\zeta} d\zeta$

$$= \frac{1}{b} \frac{f_c}{2c} \int f(\zeta) d\zeta$$

Thus, once the integral has been evaluated for a given value of b, its value for other values of b follows by proportion. This was found useful when trying the effect of different electron-density/height slopes in the vicinity of the reflection level.

(3) CALCULATIONS MADE

Ionization-density and collisional-frequency curves suitable for E-region reflections are shown in Fig. 1; (a) and (b) are different ionization distributions at low heights, and (c) and (d) are distributions at heights in the vicinity of the reflection levels. These were assessed from unpublished data due to Piggott and from rocket data.[10] They therefore represent values likely to occur commonly in practice; for instance, they lead to the following group heights not inconsistent with measurement: 102 km for the distributions (a) + (c) and (b) + (c), and 100 km for distributions (a) + (d) and (b) + (d). Varying the D-region ionization from (a) to (b) cannot give a group height change greater than 0·5 km.

Vertical- and oblique-incidence calculations were made for frequencies of 1·62 and 5·1 Mc/s, respectively. These two frequencies were not quite equivalent in the Martyn sense, but

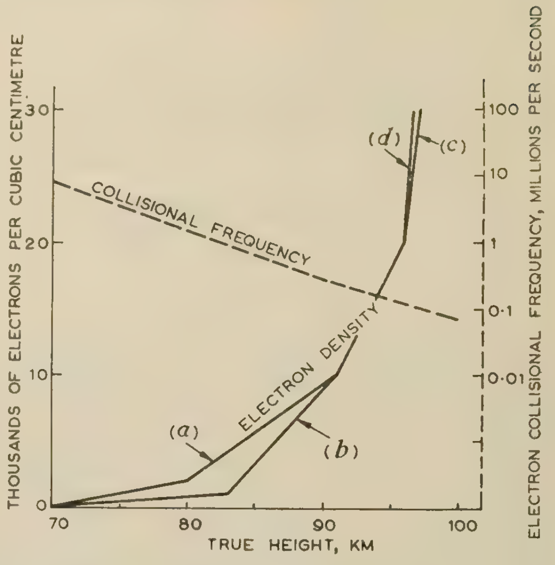


Fig. 1.—Assumed variation of electron density and collisional frequency with height.

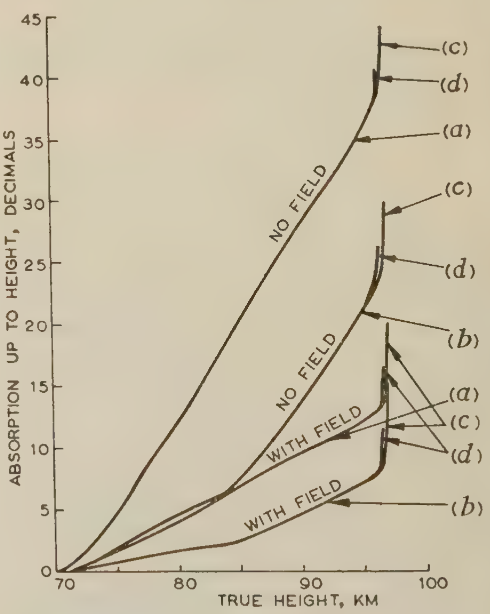


Fig. 2.—Vertical incidence; theoretical ordinary ray absorption.

Calculated from various electron-density/height distributions shown in Fig. 1. Sensitiveness to the slight change of density (c) to (d) in the vicinity of the reflection levels will be noticed.
1·62 Mc/s. Earth's field = 0·48 gauss. Dip = 69°.

re frequencies at which absorption was measured in experi-ents[4] already referred to in Section 1; they were chosen to be close to truly equivalent frequencies as conveniently possible. owever, a correction has been applied.

The angle of incidence for the (Slough–Inverness) oblique th range of 740 km was 74° (angle of elevation 16°), and the ıgnetic dip was 69°. The magnitude of the earth's magnetic d was taken to be 0·48 gauss, giving a gyro-frequency of 4 Mc/s. Preparation in the plane of the magnetic meridian s assumed.

Absorption calculated from the four different ionization con-ions in Fig. 1 is shown in Fig. 2 for vertical incidence and Fig. 3 for oblique incidence. It has been plotted in the form

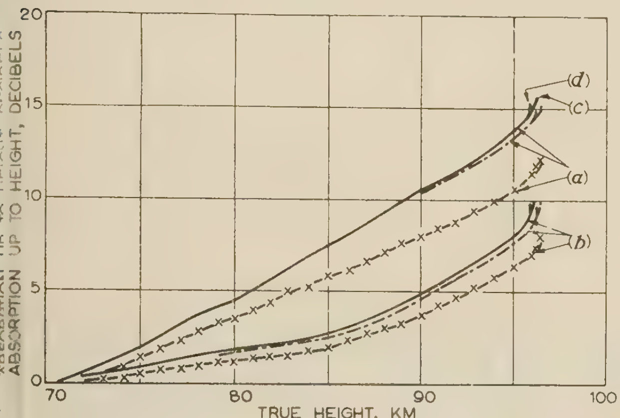


Fig. 3.—Oblique incidence; theoretical ordinary ray absorption.
alculated from various electron-density/height distributions shown in Fig. 1.
5·1 Mc/s. Earth's field = 0·48 gauss. Dip = 69°.
— · — With field, Slough to reflection level.
—×— With field, Inverness to reflection level.
——— No field, both paths.

absorption integrated up to a particular height, against that ght; the contributions at the various height levels to the total orption as the ray progresses up to the reflection level can n readily be seen. For the vertical-incidence cases the path orption is twice the final values given in Fig. 2, and the total orption between Slough and Inverness is the sum of the ugh and Inverness reflection-level half-path values given in . 3.

(4) DISCUSSION OF RESULTS

The final values of absorption of corresponding half-paths equal for all vertical-incidence cases and for the oblique es without magnetic field, but differ for the oblique cases n field. This is brought out clearly in Table 1. The fact there can be differences is a consequence of applying mag-o-ionic theory. The term 'half-path' is here used instead of going' or 'downgoing' path in order to remove any possible bts about reciprocity; the same calculation refers to a wave ng up from, say, Slough to reflection level, as to a wave ng down from reflection level to Slough, provided that the gneto-ionic components are treated separately. The mathe-ics is such that the path followed by a particular component linary or extraordinary) is the same in each case, since the active index for each component separately is independent of ction.

Each vertical-incidence curve in Fig. 2 appears to be in two s: the absorption increases more or less uniformly from m until a height of about 95 km is reached, when it rises rply to the final value. The former gives the 'non-deviative' rption and the latter the 'deviative' absorption associated with ray bending in the vicinity of the reflection level. Evidently, for our cases at any rate, the deviative absorption is considerable, even at the low frequency of 1·62 Mc/s, which is considerably below the E-region penetration frequency. In fact, for case $(b) + (c)$, with magnetic field, the deviative absorption is approximately equal to the non-deviative. It would therefore appear that the first method (Section 1) of assessing oblique absorption must in practice be applied with great care. For example, the absorption on 5·1 Mc/s calculated for oblique incidence from the $(b) + (c)$ vertical-incidence case using eqn. (1) is about 27 dB, whereas the magneto-ionic result is 17·5 dB; the 27 dB figure is too high because of the deviative absorption implicit in the vertical-incidence figure. However, some measurements have shown[3,4] that, during summer at sunspot minimum, the absorption at vertical incidence was 34 dB (with 99% confidence limits of 30 and 46 dB) at 1·62 Mc/s, and 29 dB (with 99% confidence limits of 25 and 40 dB) for the Slough–Inverness path at 5·1 Mc/s. Table 1 then shows that ionization distributions $(a) + (c)$ and $(a) + (d)$ fit these values better than the distribution $(b) + (c)$ referred to above, in which case the proportion of deviative to non-deviative absorption is smaller and errors in using eqn. (1) would not therefore be so large; but the deviative absorption cannot be made really negligible unless the rate of increase of ionization with height is perhaps 10 times greater than the maximum value here adopted, which is very unlikely for the epoch of the experiments.

On the other hand, the deviative absorption does appear to be negligible at oblique incidence (see Fig. 3), since the upward cusps in the vicinity of the reflection levels are small. It is therefore not unreasonable to suppose that eqn. (1) would then apply quite accurately. If this were true in general—and there seems to be no reason why it should not be so—the value of absorption likely to be found on one oblique path could be calculated from the value measured simultaneously on another such path traversing the same area of ionosphere, provided that sec i were not too small (greater than about 3).

The fact that eqn. (1) appears valid for the oblique case has been used to modify the values of absorption calculated at 5·1 Mc/s to those likely to occur at the more exact Martyn equivalent frequency. (It has already been mentioned in Section 3 why the more exact equivalent frequencies were not used.) The graphical method described in Section 2 had shown that the more nearly correct value was 5·7 Mc/s, with an angle of incidence of about $16\frac{1}{2}°$ and obliquity factor of 3·5, or $\cos i = 0{\cdot}29$.

The angle between the earth's magnetic field and the oblique ordinary ray at Slough is 85°, so for the half-path on the Slough side the longitudinal component, f_L, of the gyro-frequency is $1{\cdot}4 \cos 85° = 0{\cdot}012$ Mc/s, which is negligible; the propagation is, in fact, practically transverse to the earth's field, as may be seen from Fig. 3, where the value of absorption on this half-path will be seen to be only slightly less than that obtained with the earth's field absent.

On the Inverness side, with an angle of 53° between ray and field, $f_L = 0{\cdot}8$ Mc/s. It is interesting to compare this value with a quantity f_L' such that $f^2/(f + f_L')^2$ is in the inverse ratio of the field/no-field values of absorption for the half-path given in Table 1; for instance, this would be 12·2/15·7 for the field/no-field case $(a) + (c)$, giving $f_L' = 0{\cdot}7$ Mc/s.

The factor for converting the Slough reflection-level absorption from 5·1 to 5·7 Mc/s is therefore $(5{\cdot}1/5{\cdot}7)^2$, and for the Inverness reflection level is $(5{\cdot}1 + 0{\cdot}7)^2/(5{\cdot}7 + 0{\cdot}7)^2$. The results obtained, together with those for 5·1 and 1·62 Mc/s taken from Figs. 3 and 2, respectively, are given in Table 1.

In the final column the ratio of oblique (5·7 Mc/s) to vertical (1·62 Mc/s) absorption is given. When Martyn's absorption

Table 1

CALCULATIONS OF ORDINARY RAY ABSORPTION

Ionization distribution (Fig. 1)	Absorption, oblique						Absorption, vertical 1·62 Mc/s (Fig. 2)	Absorption ratio Oblique 5·7 Mc/s / Vertical 1·62 Mc/s
	5·1 Mc/s (Fig. 3)			5·7 Mc/s (corrected for)				
	(i)*	(ii)*	sum	(i)*	(ii)*	sum		
	dB	dB	dB	dB	dB	dB	dB	
(*a*) + (*c*) No field ..	15·7	15·7	31·4	12·6	12·6	25·2	88·0	0·29
Field	12·2	15·2	27·4	10·0	12·2	22·2	40·0	0·56
(*a*) + (*d*) No field ..	15·3	15·3	30·6	12·3	12·3	24·6	80·8	0·30
Field	11·7	14·8	26·5	9·6	11·8	21·4	32·6	0·66
(*b*) + (*c*) No field ..	10·0	10·0	20·0	8·0	8·0	16·0	59·2	0·27
Field	7·9	9·6	17·5	6·5	7·7	14·2	30·0	0·47
(*b*) + (*d*) No field ..	9·8	9·8	19·6	7·9	7·9	15·8	52·4	0·30
Field	7·5	9·1	16·6	6·2	7·3	13·5	22·6	0·60

* (i) Inverness reflection level.
(ii) Slough reflection level.

theorem is unquestionably valid, i.e. without field, this ratio should equal cos $i = 0{\cdot}29$; it is not always exactly so, owing to the approximation introduced by converting from 5·1 to 5·7 Mc/s on a path of constant length. However, with magnetic field present, the ratio rises to 0·66 in one instance. Taking Martyn's theorem with the cos i factor omitted (the method of calculation mentioned in Section 1), this ratio would be unity: magnetic field is therefore having the effect of making the cos i term tend to unity and is similar to partial reflections in that respect. In other words, the absorption calculated for oblique incidence from vertical-incidence measurements at the equivalent frequency would be too low using Martyn's theorem; for the instance quoted it would be in error by −12·1 dB, and by +11·2 dB if the cos i term were omitted. It is therefore not surprising that omitting the cos i term may sometimes give results of the right order in practice, since this omission seems to lead to a pessimistic value for the field strength to be expected on an oblique path, thus compensating for extra losses which are not normally taken into account. However, for an E-reflected trajectory, at least, there does not appear to be much justification for neglecting the cos i term completely. Since the effect of the earth's field is greatest at vertical incidence, as can be seen from Table 1, it is highly probable that the same applies to absorption calculated by Martyn's theorem for F-reflected trajectories, but with the possibility in practice of additional losses due to partial reflections from sporadic-E ionization. In this respect, Beynon's[2] Table 2 shows that absorption calculated (using Martyn's theorem with the cos i factor omitted) for a transatlantic path at 9·87 Mc/s was on the average about 15 dB too high during noon in winter, whilst it was 5 dB too low in summer, when more sporadic-E reflections might have been expected. There is therefore a seasonal variation of 20 dB, at least part of which might conceivably be attributed to partial reflections. On the other hand, it might also be attributable to the effect of the 'winter anomaly'[11] on the Slough vertical-incidence figures; this phenomenon causes days of high absorption to occur during the winter, so that the monthly averages from November to March are anomalously elevated. There is evidence to show that this effect is local,[12] and so winter noon observations at Slough might be considerably in error if applied to the midpoint of a long path. However, further experimental work would be required on oblique paths to see whether they, as well as those at vertical incidence, were anomalous in this respect during the winter.

(5) CONCLUSIONS

Absorption along a vertical and an oblique path reflected in the E-region has been calculated for various ionization distributions and the results compared. Millington's form of Booker's ray theory, which includes the earth's magnetic field, was used. Frequencies of 1·62 Mc/s and 5·7 Mc/s were taken at vertical and oblique incidence, respectively, since the ratio between these frequencies is the cosine of the angle of incidence for the oblique path of 740 km utilizing single-hop E-reflections; the frequencies were therefore equivalent in the sense of Martyn's theorem. The following conclusions of practical significance can be reached from these calculations:

(*a*) For vertical incidence, absorption in the vicinity of the reflection level (deviative absorption) cannot always be neglected compared with the (non-deviative) absorption taking place in the lower E, or D, region. This is especially true when the earth's magnetic field is taken into account, even though the frequency used may be considerably below the E-region critical, or penetration, frequency. This tends to invalidate the simple formula for non-deviative absorption which is often used to convert absorption measured at vertical incidence to that at oblique incidence at the same frequency.

(*b*) For the oblique case, however, deviative absorption during the process of reflection can be neglected compared with non-deviative. The non-deviative formula referred to in Section 1 should therefore be applicable when calculating absorption to be expected on one oblique path from measurements made on another, the path being split into up and down components with a value of f_L appropriate to each; for a better approximation Fig. 13 of Reference 6 (second paper) could be used, but the answer is not critically dependent on the values of f_L taken. Clearly, the obliquity of the paths must be sufficient for this to be valid; it can be inferred from the calculations that, with E-region reflections, angles of incidence greater than about 70° (elevations less than 20°) would qualify.

(*c*) The fact that better results are usually obtained in practice by omitting the cos i term in Martyn's theorem can be mainly accounted for when a full magneto-ionic treatment is used. (Martyn's theorem was not intended to be valid in the presence of the earth's magnetic field.) It does not seem necessary, therefore, to invoke partial reflections as an explanation, at least for rays reflected at the E-region. For practical purposes there does

appear to be any reason why Martyn's theorem, either with without the cos i term, should be relied on, except as a very gh guide indeed.

d) The calculation of absorption using magneto-ionic theory edious, and, it is suggested, is unjustifiable in general for ticular radio-circuit applications, since the results depend siderably on the particular distributions of electron density l collisional frequency used. However, the E-region cal- ations made in the present paper have indicated that ique-incidence absorption calculated from vertical-incidence ues may be in error and should be used with care. To nplete the work, similar calculations should be made for ne typical F-region paths. If a knowledge of absorption on a ticular oblique path is required, it seems better to estimate it m that on another such path, or from an experimental one up for the purpose. The need for relating the absorbing racteristics of the ionosphere on the two paths, however, uld remain, and vertical-incidence soundings seem useful in respect only if deviative and non-deviative absorption can properly separated. Single-hop oblique-incidence absorption surements, using pulses, conducted simultaneously in various s of the world should prove more useful than vertical- dence measurements in estimating these characteristics.

(6) ACKNOWLEDGMENTS

he work was carried out as part of the programme of the io Research Board. The paper is published by permission the Director of Radio Research of the Department of ntific and Industrial Research.

(7) REFERENCES

PIGGOTT, W. R.: Discussion on Ionospheric Absorption and Reflection of Radio Waves, *Proceedings I.E.E.*, 1954, **101**, Part III, p. 370.

(2) BEYNON, W. J. G.: 'Some Notes on the Absorption of Radio Waves reflected from the Ionosphere at Oblique Incidence', *ibid.*, Paper No. 1569 R, January, 1954 (**101**, Part III, p. 15).

(3) MEADOWS, R. W., and MOORAT, A. J. G.: 'The Effect of Fading on the Accuracy of Measurement of Ionospheric Absorption' (see page 27).

(4) MEADOWS, R. W.: 'The Attentuation of Radio Waves reflected from the E-Region of the Ionosphere' (see page 22).

(5) ALLCOCK, G. MCK.: 'Ionospheric Absorption at Vertical and Oblique Incidence', *Proceedings I.E.E.*, Paper No. 1664 R, April, 1954 (**101**, Part III, p. 360).

(6) MILLINGTON, G.: 'The Effect of the Earth's Magnetic Field on Short-Wave Communication by the Ionosphere', *ibid.*, Monograph No. 1 R, April, 1951 (**98**, Part IV, p. 1).
'Ray-Path Characteristics in the Ionosphere', *ibid.*, Monograph No. 94 R, April, 1954 (**101**, Part IV, p. 235).

(7) BOOKER, H. G.: 'Propagation of Wave-Packets incident obliquely upon a Stratified Doubly-Refracting Ionosphere', *Philosophical Transactions of the Royal Society*, 1938, **237**, No. 781, p. 411.

(8) BOOKER, H. G.: 'Applications of the Magneto-Ionic Theory to Radio Waves incident obliquely upon a Horizontally Stratified Ionosphere', *Journal of Geophysical Research*, 1949, **54**, p. 243.

(9) 'Fundamental Principles of Ionospheric Transmission', Radio Research Special Report No. 17, 1948 (Code No. 47-29-17).

(10) SEDDON, J. C., PICKAR, A. D., and JACKSON, J. E.: 'Continuous Electron Density Measurements up to 200 km', *Journal of Geophysical Research*, 1954, **59**, p. 513.

(11) APPLETON, E. V., and PIGGOTT, W. R.: 'Ionospheric Measurements through a Sunspot Cycle', *Journal of Atmospheric and Terrestrial Physics*, 1954, **5**, p. 141.

(12) BEYNON, W. J. G., and DAVIES, K.: 'Simultaneous Ionospheric Absorption Measurements at Widely Separated Stations', *ibid.*, 1954, **5**, p. 273.

621.396.677

The Institution of Electrical Engineers
Paper No. 2430 R
Jan. 1958
©

AN INVESTIGATION OF PERIODIC ROD STRUCTURES FOR YAGI AERIALS

By J. O. SPECTOR, Ph.D., Associate Member.

(*The paper was first received* 17*th May, and in revised form* 9*th August*, 1957.)

SUMMARY

The periodic structure of conducting rods, normally used for Yagi-type aerials, is investigated from the point of view of a guide for surface waves. In a series of resonator experiments it is found that a non-radiating plane surface wave of the HE_{11}-mode may be guided along the structure, and that radiation occurs only from a discontinuity. The guiding effect is found to exist only where the rod lengths are less than $\lambda/2$. The structure is thus analogous to a capacitively-loaded transmission line. The propagation coefficients are experimentally determined for structures of different rod length and spacing. These propagation coefficients are used to predict the radiation patterns of Yagi aerials in conjunction with the theory outlined in a previous paper. There is quite good agreement with experimentally-observed radiation patterns of long Yagi aerials. The side-lobe structure is explained by interference with direct radiation from the driven element caused by inefficient launching of the surface wave.

LIST OF PRINCIPAL SYMBOLS

λ_0 = Free-space wavelength of frequency f_0.
λ_g = Guided wavelength on guiding structure.
ρ, ϕ = Cylindrical co-ordinates.
θ = Angle with aerial axis.
$g(\theta)$ = Amplitude radiation pattern.
$F(\rho, \phi)$ = Amplitude distribution in aperture plane.
β, β_0 = Phase-change coefficients of guided wave and free-space wave, respectively.
Z_0, Z_1 = Characteristic wave impedances of unloaded and loaded transmission line, respectively.
d = Spacing of periodic elements.
h = Length of copper rods.
ϵ_r = Relative permittivity of supporting dielectric rod.
v_1, v_0 = Phase velocities of surface wave on periodic structure and free-space wave, respectively.
δf = Shift of resonance frequency of surface-wave resonator.
f_0 = Initial resonance frequency of surface-wave resonator, with $l = 0$.
L = Length of surface-wave resonator.
l = Length of loaded section.
K = Constant defined by eqn. (13).
$k_0 = \sqrt{(\beta^2 - \beta_0^2)}$.
A = Arbitrary constant.
E_x, E_y = Components of electric field.
$J_n(x), K_n(x)$ = Bessel functions.
$t = Z_1/Z_0$.
$u = \beta_0 \frac{h}{2} \sin\theta$.

(1) INTRODUCTION

The common Yagi aerial, as shown in Fig. 1, consists of one driven element and one or more passive, or parasitic, elements. The analysis of this structure is usually made by assuming that the driven element induces currents in the other elements, and in consequence they become radiators themselves. If these induced currents are calculated, the radiation pattern of the aerial will be given by the superposition of the individual radiation fields contributed by each element. To calculate the currents, the structure is considered equivalent to a linear network,[1] and the voltage and current relations are given by

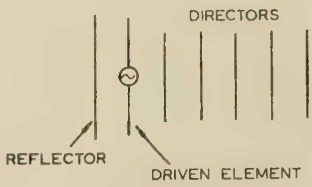


Fig. 1.—Yagi aerial.

$$[Z_{nm}]\,[I_n] = [V_n] \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where V_n and I_n are voltages and currents in the element n of the structure and Z_{nm} is the mutual impedance (or self-impedance in the case $n = m$) of the elements denoted by the subscripts. In the case of the Yagi aerial, where only one element p is driven, V_n in eqn. (1) is made zero except when $n = p$.

When eqn. (1) is applied in practice, it is immediately observed that the calculation of the coefficients Z is extremely difficult, as there are $n(n+1)/2$ such coefficients for an n-element aerial. Indeed, this approach has not yielded useful results for structures of more than a few elements.

A new approach to the problem has been suggested.[2] The periodic structure of conducting elements is considered as an artificial dielectric, along which a surface wave[3,4] is propagated. No radiation takes place from points along the length of the structure. There is, however, radiation from the discontinuity at the end of the structure, and the mechanism of radiation can be described in a very similar way to that of the dielectric rod and other end-fire structures.

A complete theoretical analysis of the Yagi-type structure in terms of a non-radiating plane surface wave would require the knowledge of the field configuration in the vicinity of the structure. By formulating continuity conditions for the fields at the boundary of the structure the propagation coefficient could be calculated, following the analysis given by Stratton[5] for plane waves on cylindrical guides. The radiation pattern may now be calculated by assuming the transverse components of the field outside the structure to be identical to those of a plane surface wave of the 'dipole mode' type. The radiation pattern of a transverse field distribution $F(\rho, \phi)$ in the aperture plane is given by[6]

$$g(\theta, X) = (1 + \cos\theta)\int_0^{\infty}\int_0^{2\pi} F(\rho, \phi)\exp\left[j\beta_0 \rho \sin\theta \cos(X - \phi)\right]\rho d\rho d\phi \quad (2)$$

where $g(\theta, X)$ is the field intensity at a point on the surface of a sphere specified by the polar angles θ and X. However, the field

Written contributions on papers published without being read at meetings are invited for consideration with a view to publication.
Dr. Spector was formerly in the Electrical Engineering Department, Imperial College of Science and Technology, London University, and is now at the Ministry of Defence, Israel.

figuration near the structure is unknown. It is therefore posed to find the propagation coefficient experimentally, and, uming that any distortion of the field near the structure reases very rapidly in the radial direction, to find the radiation tern by using the field distribution of a pure dipole-mode ve in conjunction with the aperture field method given in . (2).

The purpose of the experimental investigation of the problem to find whether the periodic structure shown in Fig. 2 could

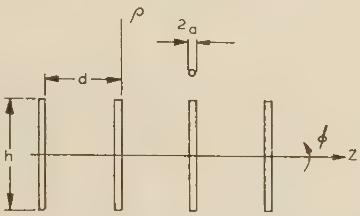


Fig. 2.—Periodic structure of conducting rods.

vide a guiding effect which is required for the propagation of on-radiating surface wave along itself, and, once the existence uch a guiding effect is established, to find the propagation fficient as a function of the structural dimensions. Finally, as proposed to find whether, using these experimental results, uantitative as well as qualitative description of the Yagi aerial ld be given, which would be consistent with measured iation patterns of Yagi aerials.

(2) ANALOGY WITH LOADED TRANSMISSION LINE

ome results of a qualitative nature can be found by using transmission-line analogy for any particular mode of propagaon. The hypothetical transmission of a homogenous plane e in the absence of the structure of conducting rods can be esented by a transmission line of characteristic impedance Z_0 phase-change coefficient $\beta_0 = 2\pi/\lambda_0$. The addition of conting rods periodically along the line of propagation may be esented in this transmission-line analogy by impedances nted across the line at regular intervals. The analysis of the odically-loaded transmission line shown in Fig. 3 presents no

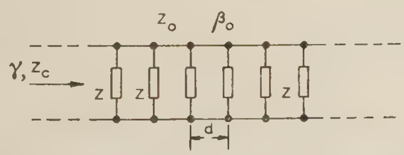


Fig. 3.—Loaded transmission line.

and β_0 are the characteristic impedance and phase-change coefficient, respectively, e unloaded line.

culties, and the propagation coefficient $\gamma = \alpha + j\beta$ of the ed line is given by[6]

$$\cosh \gamma d = \cos \frac{2\pi}{\lambda_0} d + \frac{jZ_0}{2Z} \sin \frac{2\pi}{\lambda_0} d \quad . \quad . \quad . \quad (3)$$

re Z is the impedance of the load.

it is assumed that there is no radiation from elements along structure, and if losses in the elements themselves are neged, the shunt impedance becomes a pure reactance $Z = jX$ the line will propagate unattenuated waves ($\gamma = j\beta$) whenever

$$\left| \cos \frac{2\pi}{\lambda_0} d + \frac{Z_0}{2X} \sin \frac{2\pi}{\lambda_0} d \right| \leqslant 1 \quad . \quad . \quad . \quad . \quad (4)$$

The guided wavelength λ_g will be given by

$$\cos \frac{2\pi}{\lambda_g} d = \cos \frac{2\pi}{\lambda_0} d + \frac{Z_0}{2X} \sin \frac{2\pi}{\lambda_0} d \quad . \quad . \quad . \quad (5)$$

It can be seen from eqn. (5) that the guided wavelength is a function of the shunt impedance and the spacing d.

For small spacings ($d/\lambda \ll 1$) and large shunt impedances, eqn. (5) can be rewritten

$$-\left(\frac{2\pi d}{\lambda_g}\right)^2 \simeq -\left(\frac{2\pi d}{\lambda_0}\right)^2 + \frac{Z_0}{2X} \frac{2\pi d}{\lambda_0} \quad . \quad . \quad . \quad (6)$$

which means that X must be negative in order to reduce the phase velocity ($\lambda_g < \lambda_0$), thus making the wave physically feasible. Hence it can be deduced from the transmission-line analogy that capacitive loading is required. This suggests that the shunt element may be realized in practice by a cylindrical conductor of less than the half-wave resonant length.

(3) GENERAL DESCRIPTION OF MEASURING TECHNIQUE AND APPARATUS

A general block schematic of the equipment used for the experiments is given in Fig. 4. The periodic structure was mounted in a surface-wave resonator,[7] which consisted of two

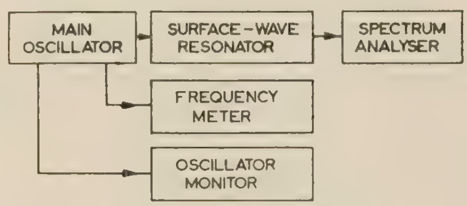


Fig. 4.—Block schematic of apparatus.

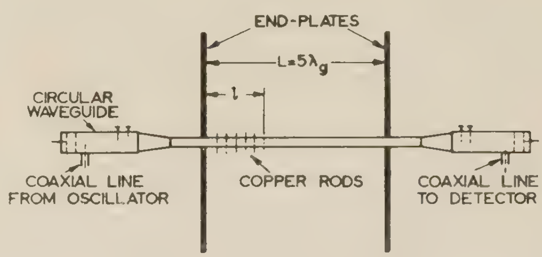


Fig. 5.—Surface-wave resonator.

aluminium plates parallel to one another and at right-angles to the axis of the structure, as shown in Fig. 5. The dimensions of the end-plates were 60 cm × 60 cm, which were large enough to contain almost all the energy of the surface-wave fields inside the resonator, for a wavelength of 10 cm used throughout the experiments. Energy was fed into the resonator through circular holes of 5 cm diameter cut in the centre of each resonator plate. The copper conductors were embedded in cylindrical holes drilled in a polystyrene rod of circular cross-section. The dielectric rod provided some initial guidance for waves along its surface. A rod of 2·54 cm diameter was used, along which the phase velocity of the HE_{11}-wave was reduced by 2×10^{-3} of its free-space value.[8]

The HE_{11} surface wave on the dielectric rod was launched by means of a circular dielectric-filled waveguide in which the dominant H_{11}-mode had been excited. The dielectric was constructed so as to continue from the end of the waveguide and

was gradually reduced to the dielectric-rod diameter by means of a taper section. Power was fed into the waveguide by means of coaxial cable and probe from a type CV238 reflex klystron.

The output section was identical to the input section. High-frequency power was transmitted by coaxial cable to a spectrum analyser, where it was rectified and presented on a cathode-ray tube in the form of a vertical deflection whose height varied according to the power input.

The frequency of oscillation was accurately measured by a heterodyne crystal-calibrated frequency meter. A small amount of power was extracted for this purpose directly from the oscillator cavity by coaxial cable.

In the absence of any copper rods, the end-plates were adjusted in such a position that a resonance could be observed on the cathode-ray tube for a frequency within the range of the oscillator valve. At the resonance frequency of the system, maximum power was obtained at the spectrum-analyser input terminals, and thus a maximum deflection on the cathode-ray-tube screen was observed. The resonance frequency was found by tuning simultaneously the cavities of the main oscillator valve and the beat-frequency oscillator of the spectrum analyser, while observing the height of the deflection on the cathode-ray tube. At the same time care was taken to operate both klystrons at the centre of their modes of oscillation, so as to avoid spurious amplitude changes, by adjusting the reflector voltages. When a resonance had been established the tuning of the beat-frequency oscillator cavity was left unchanged and its frequency-modulating voltage reduced, thus producing a large shift of the deflection along the time-base of the cathode-ray tube for a small change in main oscillator frequency. The resonance point could be located accurately on the screen, and the frequency of resonance measured by means of the frequency meter.

The system of guiding structure and end-plates is analogous to a transmission line short-circuited at both ends, for any one mode of propagation in the structure. The condition for reasonance, or natural oscillations, on such a line is that the sum of impedances 'seen' from any cross-section of the line is equal to zero.[9] The resonator partly filled to a length l with conducting rods is therefore equivalent to the section of transmission line shown in Fig. 6, where v_0 is the phase velocity along

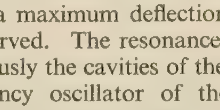


Fig. 6.—Transmission line short-circuited at both ends.

the dielectric rod and v_1 is the phase velocity along the periodic structure of length l. Taking the cross-section at the discontinuity of the line, the conditions for resonance are*

$$\tan\frac{\omega}{v_0}(L-l) + \tan\frac{\omega}{v_1}l = 0 \quad . \quad . \quad . \quad (7)$$

When no copper rods are present ($l = 0$), the frequency is denoted by ω_0, and the condition for resonance is

$$\frac{\omega_0}{v_0}L = n\pi$$

or

$$L = n\frac{\lambda_0}{2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

where n is an integer and $\omega_0/v_0 = 2\pi/\lambda_0$. This means that, at resonance, the length of the resonator must be equal to an integral number of half-wavelengths in the resonator. When conducting rods are inserted, the resonance frequency is $\omega = \omega_0 + \delta\omega$. From eqn. (7) we have, using eqn. (8) and imposing the condition that the number of half-wavelengths on the resonant line is kept constant,

$$\frac{\delta\omega}{\omega_0} = \frac{1 - v_0/v_1}{L/l + v_0/v_1 - 1} \quad . \quad . \quad . \quad . \quad (9)$$

If the change in phase velocity is small, i.e. $v_0/v_1 \simeq 1$, the change in resonance frequency is given by

$$\frac{\delta\omega}{\omega_0} = -\frac{l}{L}\left(1 - \frac{v_1}{v_0}\right) \quad . \quad . \quad . \quad . \quad (10)$$

From eqn. (10) it can be seen that, by plotting the change in resonance frequency as a function of the length occupied by the periodic structure, a straight line should be obtained. The phase velocity in the periodic structure will be given by the slope of this line. A decrease in frequency is expected for a wave which is slowed down in the structure.

The experimental procedure was as follows:

The dielectric rod was inserted in the resonator and the oscillator frequency was changed until a resonance was observed; ω_0 was then measured. The number of half-wavelengths, n, in the resonator was obtained by moving a thin conducting disc along the rod, with its axis parallel to the rod. The condition for resonance was unaffected only when the obstacle was placed at a node of the electric field. By counting the number of power-transfer maxima on the analyser screen while the obstacle was traversing the length of the resonator, the number of half-wavelengths was found. Copper rods were then added at regular intervals, starting either from the input or output end of the resonator. After each rod had been inserted, the oscillator was again tuned to resonance and the frequency was measured. This procedure was repeated until the whole length of the resonator was occupied by the periodic structure.

The amount of radiation from the structure was indicated by the change in the Q-factor of the resonator, assuming all other losses to be constant. To measure the Q-factor, the oscillator was detuned from resonance on either side until the deflection on the cathode-ray tube was half the value at resonance. The two frequencies thus found corresponded to half-amplitude points of the resonator. The bandwidth was measured either by means of the frequency meter or directly on the screen of the cathode-ray tube with the aid of the 1 Mc/s spaced markers, the first method being more accurate.

(4) DISCUSSION OF RESULTS

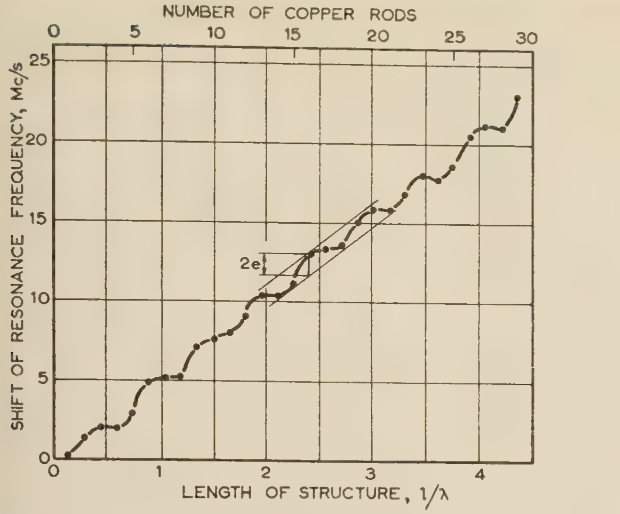
(4.1) Determination of Phase Velocities

The diameter $2a$ of the conducting rods has been kept constant at $2a = 0{\cdot}15\,\text{cm} \simeq 0{\cdot}015\lambda$, and the effects of varying the length of rod h and the spacing d have been investigated.* A large number of measurements have been carried out, a typical set being given in Fig. 7. Results are summarized in Figs. 8–10.

In the experiments carried out, a resonance could always be observed in the presence of the periodic structure of conducting rods, and so it was proved that propagation along the structure actually existed. The negative sign of δf indicated that the wave was slowed down along the structure. When the structure was formed by elements of $h/\lambda > 0{\cdot}5$, the Q-factor fell off rapidly as the first elements were inserted and no resonance could be detected. This was expected as the line was then being loaded inductively, which would make $\lambda_g > \lambda_0$.

* In eqn. (7) it is assumed that the line is loss-free and that the two parts of the line have the same characteristic impedance. The effect of a change in characteristic impedance is treated in Section 8.

* The length h is the effective length of copper rods. The length of the portion of copper rod embedded in the dielectric is multiplied by $\sqrt{\epsilon_r}$. Also, since changes of wavelength are small, the normalizing factor λ has been taken as $\lambda = 10$ cm throughout



;7.—Shift of resonance frequency as a function of structure length.
The rod length is 0·24λ and the spacing is 0·15λ.

. Fig. 7 the shift in resonance frequency, $-\delta f$, is plotted nst structure length for closely spaced structures, with = 0·15. It is seen that the measured values of δf lie on a e which oscillates with a period of $\lambda/2$ about a straight line ugh the origin. From the hypothesis of the existence of a v' wave along the periodic structure and in conjunction with equivalent circuit of Fig. 6 and eqn. (10), measured values f against l/L are expected to lie on a straight line through origin, for small reductions in phase velocity. This is, in the case in Fig. 7, apart from the small variations, periodic superimposed on this line. This oscillation may be explained ne fact that in the equivalent circuit of Fig. 6 the two sections ansmission line have been assumed to be of the same charac- tic impedance. This is, however, not the case, since the line been loaded periodically by capacitive shunt-elements. In nodified equivalent circuit Z_0 and Z_1 are the characteristic edances of the unloaded and loaded transmission lines, ectively. It is shown in Section 8 that the change in nance frequency as a function of loaded length l, for small ges of phase velocity, is given approximately by

$$-\frac{\delta f}{f_0} = t\frac{l}{L}\left(1 - \frac{v_1}{v_0}\right) + (t-1)\frac{\lambda_0}{4\pi L}\sin 4\pi\frac{l}{\lambda_0} \quad . \quad (11)$$

re $t = Z_1/Z_0$.

ne observed deviations from a straight line in Fig. 7 are ently explained by eqn. (11). If the maximum vertical ation is $2e$, the absolute value of $(t-1)$ is given by

$$|t-1| = (2e)2\pi\frac{L}{\lambda_0} \quad . \quad . \quad . \quad . \quad (12)$$

l for lines with no attenuation $(t-1)$ is real and either tive or negative. The sign of $(t-1)$ can be found by rving the phase of the oscillatory deviations in Fig. 7. This und to be negative, which again is consistent with the fact Z_1 is the characteristic impedance of a transmission line h has been shunt-loaded capacitively, so that $Z_1 < Z_0$.

ne treatment leading to eqn. (11) obviously breaks down the spacing is more than a small fraction of the wavelength. ow more reasonable to use the equivalent circuit of a smission line periodically shunt-loaded with reactive ele- ts, shown in Fig. 3. When this analysis is carried out for

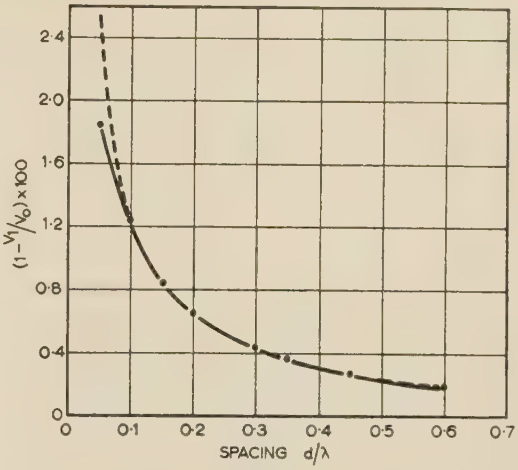


Fig. 8.—Effect of varying the spacing d/λ on the percentage reduction of phase velocity on rod structures.
The rod length is 0·24λ.
The curve of $(1 - V_1/V_0) = K\lambda/d$ is shown by the broken line, for $K = 0{\cdot}125$.

values of d/λ which are no longer small compared with the wave-length of operation, it is found that the observed deviations from a straight line are adequately explained.

In Fig. 8 the measured change in phase velocity is given as a function of spacing between the conducting rods, for rods of fixed length h. It is seen that when $d/\lambda \geqslant 0{\cdot}1$ the reduction of phase velocity is given approximately by the simple rule

$$\left(1 - \frac{v_1}{v_0}\right) = K\left(\frac{\lambda}{d}\right) \quad . \quad . \quad . \quad . \quad . \quad (13)$$

where K is a function of h only.

The experimental results are eventually summarized in Fig. 9, where the change in phase velocity is given as a function of rod length h/λ, with d/λ as parameter.

(4.2) Radiation from Yagi Aerials

In the experimental results of the previous Section it had been shown that a surface wave was established in the resonator which propagated along the structure with reduced phase velocity. The fact that a sharp resonance could be observed, even when the rod structure was long in terms of wavelengths, would itself lead to the conclusion that no large amount of power was being radiated from points along the line. However, some more experimental evidence was sought on this point, because it was of some importance to verify the proposition that any radiation from Yagi-type periodic structures occurred not at points along the structure but at the discontinuity at the end of the structure only.[2] A series of experiments was therefore carried out in which the conducting rods were, as in the previously described experiments, inserted in the dielectric guide, but attention was concentrated on the change in the Q-factor of the resonator. It was argued that if radiation occurred from the length of the structure, the Q-factor would decrease continually as l/L, the length of the loaded section, increased. If, on the other hand, radiation was confined to the discontinuity, the Q-factor should drop initially, then remain constant and finally increase again when $l/L \to 1$, as the discontinuity would then disappear.

Measurements were made on closely-spaced structures of $d/\lambda = 0{\cdot}1$ and four different rod lengths h/λ. The absolute values of Q-factor obtained were of the order of several hundred, the main source of losses being the power lost through the

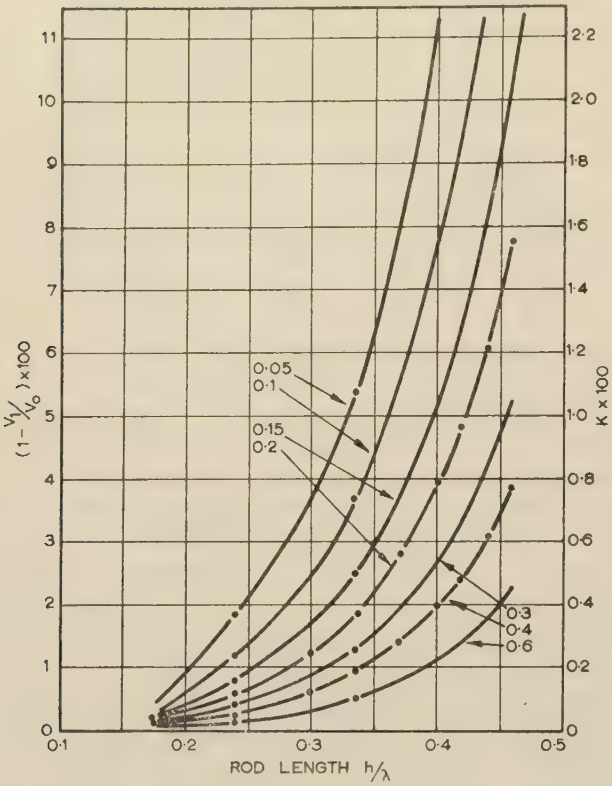


Fig. 9.—Percentage reduction of phase velocity on periodic rod structures.

The parameter is the spacing d/λ.
The right-hand scale gives the constant K of eqn. (13).

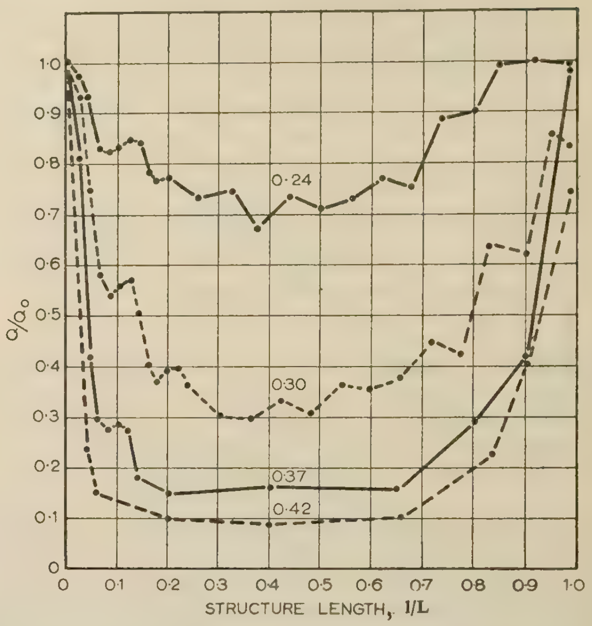


Fig. 10.—Variation of resonator Q-factor with length of rod structure.

The parameter is the rod length h/λ.
The spacing is kept constant at $d/\lambda = 0\cdot1$.
Q_0 is the value of the Q-factor when no rods are present.

coupling holes and the effect of finite dimensions of the end-plates, with the result that a small part of the field extended outside the resonator. These effects were of constant magnitude for any particular structure, and the relative Q-factors, i.e. the actually measured Q-factor divided by the initial value when no rods were present, are shown in Fig. 10. It is evident from Fig. 10 that the Q-factor falls from its initial value for the unloaded dielectric rod line to a lower value which depends on the length of the conducting rods. The greater the difference in phase velocity between the periodic structure and the dielectric rod line, the more pronounced is the decrease in the Q-factor. The decreased Q-factor remains almost constant as l/L is increased, indicating that the amount of loss does not increase. As the periodic structure occupies the whole length L of the resonator, the Q-factor increases again until, at last, for $l/L = 1$ it is almost restored to its initial value. Results given in Fig. 10 therefore provide quite conclusive evidence that, unless there is a discontinuity, no radiation takes place from a Yagi-type periodic structure.

It is now suggested that the radiation mechanism of Yagi aerials can be explained in terms of a surface wave propagating along the aerial and a radiating aperture at the end of the aerial structure. By means of a driven element, a surface wave of the dipole-mode type is excited which propagates without radiation along the aerial, and the radiation pattern can be derived from the field distribution in the terminal aperture plane, using eqn. (2). Approximations may now be made on the assumption that the phase-change coefficient, β, of the wave on the structure is almost equal to the free-space phase-change coefficient β_0. Thus, in terms of Cartesian co-ordinates, the field components in a transverse plane outside a cylinder $\rho = h/2$ bounding the rod structure are given by[2]

$$E_x = A\mathrm{K}_0(k_0\rho) \quad . \quad . \quad . \quad . \quad . \quad (14)$$

$$E_y = 0 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (15)$$

where

$$k_0^2 = \beta^2 - \beta_0^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad (16)$$

The exact field distribution for $\rho \leqslant h/2$ is not known, but, to the extent of the above-mentioned approximations, the contribution of this part of the radiating aperture can be neglected, since, for $\lambda_g/\lambda_0 \simeq 1$, a far greater part of the energy of the dipole wave is transmitted outside the structure. Substitution of eqn. (14) into eqn. (2) gives

$$g(\theta, X) = 2\pi A(1 + \cos\theta)\int_{h/2}^{\infty} \rho\mathrm{K}_0(k_0\rho)\mathrm{J}_0(\beta_0\rho\sin\theta)d\rho \quad (17)$$

which shows that the radiation pattern is independent of X. This is consistent with observed radiation patterns of long Yagi aerials, which are very similar in the plane of polarization and perpendicular to it.[10]

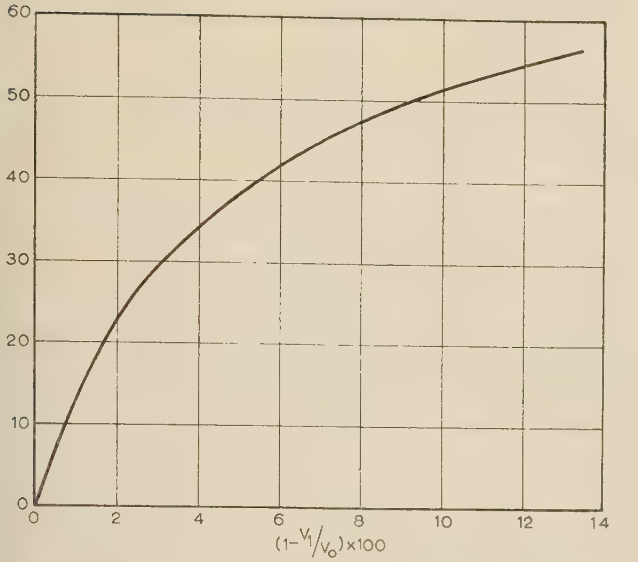
The integral of eqn. (17) is of the Lommel type, and integration yields

$$g(\theta) = \tfrac{1}{2}\pi Ah^2(1 + \cos\theta)\,\frac{\frac{k_0h}{2}\mathrm{K}_1\left(\frac{k_0h}{2}\right)\mathrm{J}_0(u) - u\mathrm{K}_0\left(\frac{k_0h}{2}\right)\mathrm{J}_1(u)}{\left(\frac{k_0^2h^2}{4} + u^2\right)} \quad (18)$$

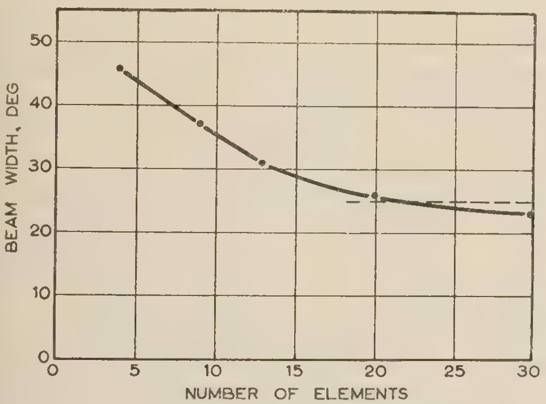
where

$$u = \beta_0\frac{h}{2}\sin\theta \quad . \quad . \quad . \quad . \quad . \quad . \quad (19)$$

The expression in eqn. (18) can be readily evaluated with the help of Tables. This can be carried out for any given reduction in phase velocity, and the resulting half-amplitude beam width is shown in Fig. 11. From this figure and the phase velocities



. 11.—Theoretical dependence of half-amplitude beam width on the percentage difference between free-space and dipole-wave phase-velocity.



. 12.—Experimental beam width of Yagi aerials from Reference 11.
ne rod length is 0·4 λ and the spacing is 0·34 λ.
ne beam width of long Yagi aerials calculated by the method outlined in the er is shown by the broken line.

ained from the resonator experiments, the beam width of a g Yagi array can be predicted.

set of experimental results for long Yagi aerials is shown Fig. 12, where the aerial length is increased and all other ameters are kept constant.[11] The beam width calculated by new approach is also indicated. It appears to tend to about for large numbers of director elements. The corresponding oretical figure, as predicted from Fig. 11, using the experi- ntally determined dipole-mode wavelength for a structure ing the same rod lengths and spacings as the Yagi array, is . Agreement is quite good, although one would expect the oretical beam width to be smaller than the experimental.

t should be noted, however, that this analysis is strictly valid y when all the power fed to the aerial is converted into a ace wave, and will apply only to aerials which are long in s of wavelengths. In practice, Yagi aerials are quite short, some part of the input power is radiated directly from the ver element. The direct radiation from the driven element either increase or decrease the beam width.[2] The super- tion of the two radiation patterns will contribute to the be structure.

An aerial with short elements would have a larger radiating aperture, as the field of the surface wave becomes more wide-spread about the guiding structure, resulting in a sharper beam, but at the same time one would expect a lower launching efficiency of the surface wave,[12] resulting in a greater amount of power being radiated directly from the driven dipole, which has a low directivity. The optimum construction of the aerial for best overall performance would have to be found experimentally.

(5) CONCLUSIONS

It has been shown that a non-radiating plane surface wave is supported by a Yagi-type periodic structure of conducting rods, and that radiation from the structure occurs only from a discontinuity. The rod structure has been found to be analogous to a capacitively loaded transmission line for rod lengths less than $\lambda/2$. The dipole wave has been excited on the structure, and in a series of resonator experiments the propagation coefficients have been experimentally determined for structures of different dimensions. With the aid of the aperture field approach the beam width can be predicted when these structures are used as aerials. Results are found to be in good agreement with experimentally observed radiation patterns for long Yagi aerials.

(6) ACKNOWLEDGMENTS

The work described in the paper was carried out at the Electrical Engineering Department of the Imperial College of Science and Technology, and the author is indebted to Dr. Willis Jackson for the facilities provided. Many thanks are due to Dr. J. Brown for helpful discussions of the problem. The author wishes to acknowledge the award of a British Council Scholarship and a grant from the Technion Society of Great Britain.

(7) REFERENCES

(1) Pidduck, F. B.: 'Currents in Aerials and High Frequency Networks' (Oxford University Press, 1949).
(2) Brown, J., and Spector, J. O.: 'The Radiating Properties of End-Fire Aerials', *Proceedings I.E.E.*, Paper No. 2216 R, January, 1957 (**104 B**, p. 27).
(3) Barlow, H. M., and Cullen, A. L.: 'Surface Waves', *ibid.*, Paper No. 1482 R, April, 1953 (**100, Part III**, p. 329).
(4) Brown, J.: 'The Types of Wave which may exist near a Guiding Surface', *ibid.*, Paper No. 1567 R, November, 1953 (**100**, Part III, p. 363).
(5) Stratton, J. A.: 'Electromagnetic Theory' (McGraw-Hill, 1941).
(6) Silver, S.: 'Microwave Antenna Theory and Design' (McGraw-Hill, 1949).
(7) Barlow, H. M., and Karbowiak, A. E.: 'An Investigation of the Characteristics of Cylindrical Surface Waves', *Proceedings I.E.E.*, Paper No. 1462 R, April, 1953 (**100**, Part III, p. 321).
(8) Chandler, C. H.: 'An Investigation of Dielectric Rod as Waveguide', *Journal of Applied Physics*, 1949, **20**, p. 1188.
(9) Schelkunoff, S. A.: 'Electromagnetic Waves' (Van Nostrand, 1943).
(10) Alred, R. V.: 'Experiments with Yagi Aerials at 600 Mc/s', *Journal I.E.E.*, 1946, **93**, Part IIIA, p. 1490.
(11) Fishenden, R. M., and Wiblin, E. R.: 'Design of Yagi Aerials', *Proceedings I.E.E.*, Paper No. 731 R, January, 1949 (**96**, Part III, p. 5).
(12) Cullen, A. L.: 'The Excitation of Plane Surface Waves', *ibid.*, Monograph No. 93 R, February, 1954 (**101, Part IV**, p. 225).

(8) APPENDIX: The Resonant Frequency of a Two-Section Transmission Line Short-Circuited at Both Ends

Let Z_0 and Z_1 be the characteristic impedances of the two sections of short-circuited transmission line shown in Fig. 6. The condition for resonance at a frequency ω is given by

$$Z_0 \tan \frac{\omega}{v_0}(L - l) + Z_1 \tan \frac{\omega}{v_1} l = 0 \quad . \quad . \quad (20)$$

where v_1 is the phase velocity on the section of impedance Z_1 and length l, and v_0 is the phase velocity on the section of impedance Z_0 and length $(L - l)$. For a small change in resonance frequency we have

$$\omega = \omega_0 + \delta\omega \quad . \quad . \quad . \quad . \quad . \quad . \quad (21)$$

where the frequency of resonance for $l = 0$ is denoted by ω_0 and is given by $\omega_0/v_0 L = n\pi$. If now the value of $\delta\omega/\omega$ is made sufficiently small, substitution of eqn. (21) into eqn. (20) yields

$$\frac{\delta\omega}{v_0}(L - l) \sec^2 \frac{\omega_0}{v_0} l - \tan \frac{\omega_0}{v_0} l + t\left(\frac{\delta\omega}{v_1} l \sec^2 \frac{\omega_0}{v_1} l + \tan \frac{\omega_0}{v_1} l\right) = 0 \quad (22)$$

where $t = Z_1/Z_0$. Let it be assumed that the phase velocities of the two sections differ by a small amount only,

$$v_1 = v_0 + \delta v \quad . \quad . \quad . \quad . \quad . \quad . \quad (23)$$

Substituting eqn. (23) into eqn. (22) and neglecting second-order infinitesimals, we have

$$-\frac{\delta\omega}{\omega_0} = \frac{tl\left(1 - \dfrac{v_1}{v_0}\right) + \dfrac{v_0}{\omega_0}\left(\dfrac{t-1}{2}\right) \sin 2\dfrac{\omega_0}{v_0} l}{L\left[1 - \dfrac{l}{L}\left(1 - \dfrac{v_0}{v_1} t\right)\right]} \quad . \quad (24)$$

Assuming the value of the impedance ratio t to be near unity, and also writing

$$\frac{\omega_0}{v_0} = \frac{2\pi}{\lambda_0}$$

we have

$$-\frac{\delta\omega}{\omega_0} = t\frac{l}{L}\left(1 - \frac{v_1}{v_0}\right) + (t - 1)\frac{\lambda_0}{4\pi L} \sin 4\pi \frac{l}{\lambda_0} \quad . \quad (25)$$

which becomes identical with eqn. (10) on making $t = 1$.

621.382.333:621.317.61

The Institution of Electrical Engineers
Paper No. 2438 R
Jan. 1958

SOME MEASUREMENTS ON COMMERCIAL TRANSISTORS AND THEIR RELATION TO THEORY

By F. J. HYDE, M.Sc., Associate Member.

(*The paper was first received* 13*th February, and in revised form* 15*th August*, 1957.)

SUMMARY

The effective lifetimes of minority carriers in the bases of five types of transistor have been measured under both steady-state and transient conditions as functions of emitter current. For alloy transistors, good agreement is obtained by the two methods in an overlap range of emitter current, within which both methods are valid. The lower emitter efficiency of the surface-carrier transistor prevents a direct comparison from being made in its case.

For three types the effective diffusion constant apparently varies with emitter current over a range somewhat greater than that predicted by existing theory.

The variation of base-to-collector current gain with emitter current is discussed in terms of the separate variations of effective lifetime and diffusion constant.

LIST OF PRINCIPAL SYMBOLS

A = Emitter area.
A_s = Area of base surface surrounding emitter that is effective for recombination at low injection level.
α_{cb}, α_{cb0} = Short-circuit internal current gain, and its zero-frequency value, between base and collector.
α_{ced} = Short-circuit internal diffusion-current gain between emitter and collector.
C_E, C_C = Emitter and collector depletion-layer capacitance.
D_n, D_p = Diffusion constant for electrons and holes, respectively.
e = Electronic charge.
$f_{\alpha d}$ = Cut-off frequency of the internal short-circuit diffusion-current gain between emitter and collector.
G = Emitter input hole conductance of the internal transistor at zero frequency.
I_E = Emitter current.
k = Boltzmann's constant.
l_n = Diffusion length of electrons in p-type emitter.
l_p = Diffusion length of holes in n-type base.
μ_p = Mobility of holes.
n_p = Equilibrium electron density in p-type emitter region for zero emitter voltage.
N_a, N_d = Acceptor and donor densities in the base.
p_n = Equilibrium hole density in n-type base.
T = Absolute temperature.
τ_p = Bulk lifetime of holes in n-type base for low injection level.
τ_{pe} = Effective device lifetime for low injection level.
v = Surface recombination velocity for low injection level.
V_E = Applied emitter-to-base p.d.
V_C = Applied collector-to-base p.d.
w = Transistor base width.

α'_{cb}, D'_p, l'_p, n'_p, τ'_p, τ'_{pe}, v' and w' are effective values of the parameters defined above for arbitrary injection levels.

(1) INTRODUCTION

One-dimensional diffusion theory has been used in the literature to predict the performance of an 'ideal' transistor, i.e. one in which the emitter-base and collector-base interfaces are planar, and whose base surface plays no part in controlling the diffusion process. Despite the fact that commercial transistors are non-ideal in the above sense, it is current practice, when formulating equivalent circuits for such transistors, to assume that one-dimensional theory does apply. In the present work an attempt has been made to show how closely an analysis of measurements of current gain can be made on the basis of the one-dimensional model. In this respect, particular reference must be made to two parameters, namely the lifetime and the diffusion constant of minority carriers in the base.

(1.1) Life-Time of Minority Carriers

The recombination of injected minority carriers in the base regions of transistors cannot be rigorously described by a single decay constant (or lifetime), because recombination is controlled both by the surface[1, 2] and by the bulk material of the base. It has been shown by Moore and Pankove[3] and by Stripp and Moore[4] that the geometry of transistors and the surface treatment of the base exert a marked influence on the proportions of carriers lost through surface and volume recombination respectively. Despite this, there is much evidence that the experimental frequency-dependence of the a.c. parameters of practical transistors, under normal conditions of operation, are in close agreement with the theoretical frequency-dependence suggested when an effective lifetime, τ_{pe}, replaces volume lifetime, τ_p, wherever it occurs in the one-dimensional formulae.[5–7]

(1.2) Effective Diffusion Constant

An electric field* is developed in the base when minority carriers are injected into it,[8–10] and this field, which increases with injection level, must be taken into account in the one-dimensional transport equations. An exact solution to these equations for all levels of injection is difficult to obtain, but Rittner[9] has shown that the solution has the same form at both low and high levels. The usual low-level formulae can be used directly at high levels, provided that the low-level diffusion constant for holes (D_p for p–n–p transistors) is replaced by $2D_p$ wherever it occurs (including its hidden presence in the diffusion length) and that n_p is multiplied by

$$1 + (2p_n/N_d) \exp (eV_E/kT)$$

It may be assumed that, in the transition region between low and high levels of injection, there is a gradual change of effective diffusion constant, D'_p, from D_p to $2D_p$.

(1.3) Base-to-Collector Current Gain

Webster[8] considered the theoretical dependence of the base-to-collector current gain at zero frequency, α_{cb0}, on emitter

* The injection of minority carriers into the base is accompanied by an equivalent increase of majority carriers, which are drawn in from the base lead. The flow of these majority carriers increases the majority-carrier current from base to emitter and sets up in the base an electric field whose sense is such as to hasten the transport of minority carriers across it.

Written communications on papers published without being read at meetings are invited for consideration with a view to publication.
This paper is an official communication from the Radio Research Station, Department of Scientific and Industrial Research.

current for alloy-type transistors. Rittner[9] and Misawa[10] carried out a more rigorous analysis of the problem, as a result of which the following expression arises for α_{cb0}*

$$\alpha_{cb0} = \left(\frac{1}{2}\frac{w^2}{D_p'\tau_p'} + \frac{D_n n_p' w}{D_p' p_n l_n} + \frac{wv'A_s'}{D_p' A}\right)^{-1} \quad . \quad . \quad (1)$$

If, instead, the transistor performance is evaluated in terms of an effective minority-carrier lifetime, τ_{pe}', the expression for α_{cb0} is

$$\alpha_{cb0} = \left(\frac{1}{2}\frac{w^2}{D_p'\tau_{pe}'} + \frac{D_n n_p' w}{D_p' p_n l_n}\right)^{-1} \quad . \quad . \quad . \quad . \quad (2)$$

In these equations, and throughout the remainder of the paper, primes appear on those symbols which are, or may possibly be, dependent on the level of injection to an appreciable extent. Symbols without primes refer to the low-level values of the parameters. The use of

$$n_p' = n_p[1 + (2p_n/N_d)\exp(eV_E/kT)]$$

as discussed in Section 1.2, does not imply that n_p' varies physically in this manner, but it is a convenient algebraic relationship useful in explaining the overall change of emitter efficiency with injection level.

Apart from the variation of D_p' and n_p' discussed earlier, possible variables are A_s', v' and τ_p'. There appears to be no published information concerning variation of A_s', although it would be expected to increase approximately in proportion to the diffusion length $l_p \equiv (D_p'\tau_p')^{1/2}$ for holes in the base. Armstrong *et al.*[11] state that there is no evidence for variation of v' at high levels of injection; these authors have also shown that, for germanium of 1·5 ohm-cm resistivity, the bulk lifetime is invariant at both low and high levels of injection, but increases by something like 2 : 1 in a transition range between low and high levels.

As a result of these considerations it is seen that the first term of eqn. (1) will fall with increasing injection level, the second term will probably first fall and then rise for a well-designed transistor, while the third term will probably change only slightly. Because the third term is dominant at low levels of injection and the second at high levels (the first term is usually insignificant), α_{cb0} will first rise and then fall continuously. It is germane to this theoretical treatment, therefore, that unless v' and/or A_s' decrease markedly with increasing current, the ratio of the peak value of α_{cb0} to its zero-current value cannot exceed 2 : 1.

An extension of this theory to a.c. conditions proposed by Lo *et al.*[12] is discussed in Section 8.1, as is the alternative approach in terms of an effective minority carrier lifetime.

(2) EXPERIMENTAL WORK

The present experiments have been designed (*a*) to provide a comparison of the effective lifetimes of minority carriers in the base regions of commercial transistors under both transient and steady-state conditions; (*b*) to measure the variation of the effective diffusion constant of minority carriers in the base with emitter current; and (*c*) to study the effect of these two parameters on the measured base-to-collector current gain of the same transistors. All measurements were made using transistors of *p–n–p* configuration, at room temperature.

(2.1) Transient Measurement of Lifetime

The basic method was due originally to Gossick[13] and to Lederhandler and Giacoletto.[14] A pulse of current is fed between the emitter and the base in the forward direction, with the collector electrically floating. A high impedance is presented in the emitter circuit to the flow of excess minority carriers back from the base across the emitter junction after the forward pulse of current has ceased, a biased thermionic diode being used for this purpose.[14] The resulting emitter-to-base open-circuit p.d. is observed, using a high-input-impedance high-frequency oscillograph.

It has been shown that, provided that the emitter-to-base p.d. at the end of the pulse of current, $V_E(0)$, greatly exceeds kT/e, and that the level of injection is not too high, a simple relationship exists between the open-circuit emitter-to-base p.d., $V_E(t)$, and the effective lifetime of minority carriers in the base, namely

$$V_E(t) = V_E(0) - kTt/e\tau_{pe}' \quad . \quad . \quad . \quad . \quad (3)$$

This equation is derived on the basis of an assumed exponential decay of minority carriers, at the base-emitter interface, on cessation of the current pulse. Diffusion effects inside the base can modify this dependence, as has been shown by North (in an Appendix to the paper by Lederhandler and Giacoletto[14]), and by Henderson and Tillman.[15] Nevertheless, over a limited range of current an experimental decay having the form of eqn. (3) was actually observed.

Recurrent pulses derived from a constant-current source are used in practice. In the present work the peak emitter current, which is the independent variable, was determined by measuring the mean direct current flowing in the emitter circuit and multiplying this by an appropriate form factor based on the observed pulse/space ratio.

(2.2) Steady-State Measurements

Measurements were made of the real and imaginary parts of the internal short-circuit current gain of the transistors under test, using the bridge described by Evans[7] and illustrated in

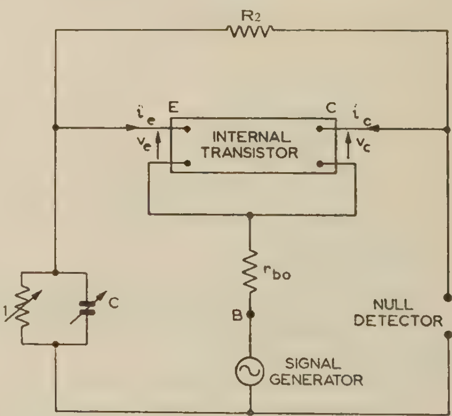


Fig. 1.—A.C. bridge used for measurement of real and imaginary parts of α_{cb} at low frequency.
The 'internal' transistor is defined by the parameters in eqn. (13).

Fig. 1. It is shown in Section 8.1 that the condition for balance of the bridge is

$$\frac{R_2}{R_1}(1 + j\omega CR_1) \simeq -\frac{(y_{11} + y_{21})}{y_{21}} = \frac{1}{\alpha_{cb}} \quad . \quad . \quad (4)$$

Here the admittance parameters y_{11} and y_{21}, which are defined by eqns. (11) and (13) relate to the 'internal transistor', i.e. the practical transistor less the ohmic base resistance. C and R_1 are variable, while R_2 is fixed at 10·32 ohms; for one transistor, C was as low as 150 pF at balance, so that care was taken to include the effect of stray capacitances.

The effect of surface recombination may be taken into account

* Rittner[9] has an additional factor of ½ multiplying the third term. This arises because he assumes that the excess minority-carrier density has fallen to zero at the extremity of the recombination surface, A_s', remote from the emitter, whereas Webster assumes that the excess minority-carrier density is uniform at its maximum value over the whole of A_s'. Because A_s' is not well defined, the algebraic distinction is unimportant.

her in the form of an effective lifetime or by using the extension oposed by Lo *et al.*[12] to Webster's[8] zero-frequency theory: is extension amounts to considering a separate 'surface-recombination admittance' which varies with frequency as $(1 + j\omega\tau_p')^{1/2}$. ne low-frequency approximations to eqn. (4), which are appropriate for the measurement frequencies used (between 1 and c/s for the alloy-type transistors and 48 kc/s for the surface-rrier transistor), are derived in Section 8.1.2 for these two parate methods of approach, where it is shown that the proposal Lo *et al.* is not in accord with the results reported by Evans[7] r a type TA153 transistor.

In the present work, therefore, the experimental results will interpreted in terms of an effective lifetime for minority rriers. Three important relationships given by eqns. (17)–(19) nich represent the low-frequency balance conditions of the idge are repeated here, namely

$$CR_1\left(1 - \frac{1}{CR_2}\frac{C_E}{G}\right) = \frac{w^2}{2D_p'}\left(\frac{w^2}{2D_p'\tau_{pe}'} + \frac{D_n n_p' w}{D_p' p_n l_n}\right)^{-1} \quad (5)$$

$$CR_2 - \frac{C_E}{G} = \frac{w^2}{2D_p'} \quad (6)$$

$$\frac{R_1}{R_2} = \left(\frac{1}{2}\frac{w^2}{D_p'\tau_{pe}'} + \frac{D_n n_p' w}{D_p' p_n l_n}\right)^{-1} \quad (7)$$

ne evaluation of C_E/G, to enable correction to be made to the t-hand sides of eqns. (5) and (6), is discussed in Section 8.2.

2.1) Discussion of Bridge Balance Conditions.

qn. (5).

Provided that $D_n n_p' w/D_p' p_n l_n \ll w^2/2D_p'\tau_{pe}'$, the right-hand side eqn. (5) is effectively equal to τ_{pe}'. This inequality is expected hold at low currents for most present-day alloy-type transistors, it not necessarily for the surface-barrier types.[16] The inequality expected to become weaker as emitter current increases, cause of the relative variations of n_p' and τ_{pe}' that may be pected (see Section 1.3).

It is difficult to make numerical allowance for the 'emitter-ficiency' term $D_n n_p' w/D_p' p_n l_n$ at high levels of injection, because lack of knowledge concerning p_n and l_n. It may be concluded, wever, that as emitter current increases and emitter efficiency lls the right-hand side of the equation will fall below τ_{pe}' to increasing extent.

qn. (6).

On the basis of existing one-dimensional theory already scussed, an increase of 2 : 1 in D_p' is expected between low and gh levels of injection, so that the right-hand side of eqn. (6) is pected to decrease in the same ratio between asymptotic limits low and high currents. From the asymptotic high-level value CR_2 it is possible to determine w, because under these conions $D_p' = 2D_p$ according to one-dimensional theory and also $_E/G$ is negligible, so that [Evans's[7] eqn. (12)]

$$w^2 = 4D_p CR_2 \quad (8)$$

qn. (7).

The relative contributions of the two terms on the right-hand de to the variation of α_{cb0} with emitter current will be deterned by the variation of n_p' and τ_{pe}' as for $CR_1(1 - C_E/CGR_2)$ scussed above.

(2.3) Interpretation of Effective Lifetime

The significance of the effective lifetime measured by the two ethods described has to be considered in the light of two ndamental differences concerning the minority carriers in the se under transient and steady-state a.c. conditions.

The first is that, in the transient case, the minority-carrier density at the base-collector interface increases above the thermal-equilibrium density,* whereas when the transistor is biased normally for steady-state a.c. operation the minority-carrier density at this interface approaches zero. The second is that an electrically-floating collector is not necessarily equivalent to a reverse-biased collector, when considered as a recombination surface.[17]

Because of these considerations, the effective lifetimes measured by transient and steady-state methods might be expected to differ somewhat. For alloy-type transistors, however, it has been shown[3] that recombination is determined mainly by a narrow surface belt surrounding the emitter, where the minority carrier density will be negligibly different in the two cases, so that no wide divergence of transient and steady-state effective lifetimes is likely.

(3) EXPERIMENTAL RESULTS

Measurements were made on germanium *p–n–p* alloy transistors of types TS2, CK760, CK762 and EW59, and on germanium surface-barrier transistors of type SB100. All transistors conformed to the manufacturer's specifications and gave repeatable results. Measurements of the d.c. characteristics of the reverse-biased emitter and collector junctions showed that the leakage paths were ohmic and had resistances greater than 1 megohm. Results are presented in detail for one transistor of each type. Two transistors of each alloy-type and three surface-barrier transistors were investigated, and in each case the results were closely comparable.

In each case graphs of CR_1, CR_2 and R_1/R_2 are plotted as functions of emitter current in Fig. 2. The experimental points from the bridge measurements are shown as dots and are joined by solid lines. Also shown are the corrected curves in which allowance has been made for C_E/G. These are shown as dot-dash curves, and represent respectively the variation of $CR_1(1 - C_E/GCR_2)$ and $CR_2 - C_E/G$.

As is explained in Section 8.2, an absolute correction for C_E/G is not possible, because of difficulty in determining C_E. For this reason the deductions that may be made from the corrected data in Fig. 2 must be regarded, to some extent, as being tentative.

The values of effective lifetime measured by the transient method are shown as crosses on the same graphs as CR_1. The crosses are joined by solid lines for that range of current in which the transient decay curves were linear and by broken lines at the extremes of current at which slight departure from linearity was manifest. When such departures became gross, no readings were taken.

The range of emitter current extended from 10 μA to the maximum working current quoted by the manufacturers. Because of variation of emitter area from one type to another, the injection level is not the same in each case for any given emitter current. An idea of the relative emitter areas may be obtained by reference to the emitter depletion-layer capacitance data given in Section 8.2. (The emitter capacitance is directly proportional to emitter area, but only to the square root of the conductivity of the base region.)

(4) DISCUSSION OF RESULTS

(4.1) Effective Lifetime

For all four alloy-type transistors there is good agreement between the effective lifetimes measured by the two different

* Gossick[13] has suggested that the density of minority carriers in the base at the base-collector interface will be virtually equal to that at the base-emitter interface, apart from a small reduction arising from recombination between the emitter and the collector. In consequence, he suggested that $V_C(0) \simeq V_E(0)$. The minority-carrier density at the collector may, however, be further reduced because of the 'feed-in feed-out' effect described by Moore and Webster[17] and additionally because the collector area is usually greater than the emitter area. Measured values of the ratio $V_C(0)/V_E(0)$ ranged from 0·6 to 0·99 in the present experiments.

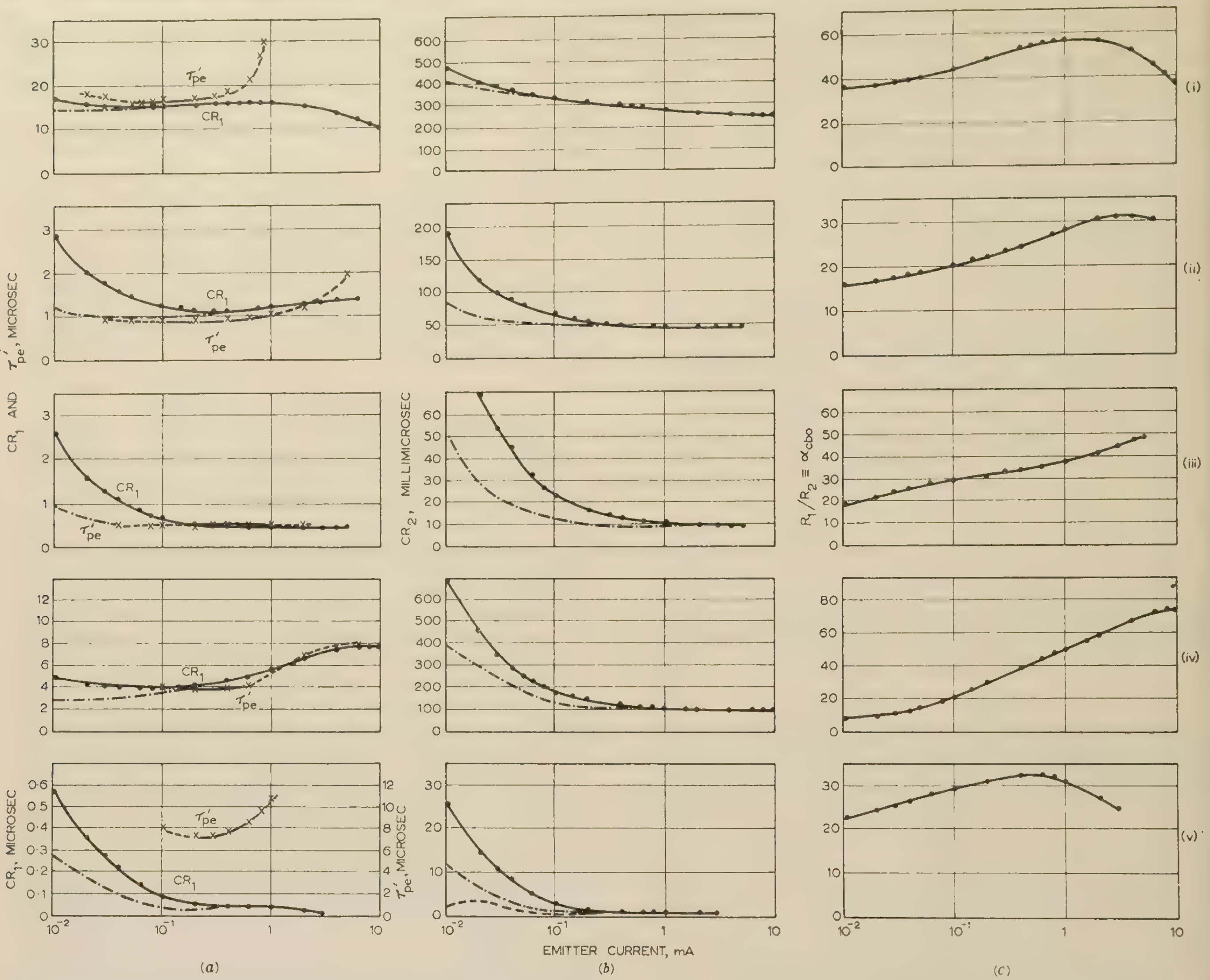


Fig. 2.—Variation of (*a*) CR_1 and τ'_{pe}, (*b*) CR_2 and (*c*) R_1/R_2 with emitter current for several types of transistor at 21° C.
(i) TS2. (ii) CK760. (iii) CK762. (iv) EW59. (v) SB100.

methods in that range of current for which both methods are valid. From the steady-state data it is evident that for two of these transistor types, namely the TS2 and EW59, there is a tendency for the effective lifetime to increase, from what is probably an asymptotic low-level value, as the level of injection is raised. For the types CK760 and CK762, however, the lifetime first decreases with increasing injection level. At still higher levels a subsequent rise is observed for the type CK760, while one may be inferred for the type CK762 if a lowering effect on CR_1 of falling emitter efficiency is assumed.

The rise in effective lifetime at high levels of injection can be interpreted in terms of the general picture presented by Webster[8] in discussing the variation of α_{cb0} with emitter current, as follows:

The proportional loss of minority carriers to the surface around the emitter decreases with increasing emitter current, because of the growth of an electric field in the base, which effectively transports away from the surface some minority carriers that would otherwise have recombined there. In consequence, the recombination of these 'reclaimed' carriers is governed by the bulk lifetime of the base material, which is much greater than the surface or effective lifetime. As was mentioned in Section 1.3, it has been shown by Armstrong *et al.*[11] that, for some germanium at least, the bulk lifetime, τ'_p, itself increases with injection level. Such an increase would contribute to the observed rise in τ'_{pe}, but only in a small way, because the effective lifetimes measured are much smaller than those of the bulk material of the base. There is an opposing effect, which should also be considered. With the increase of emitter current, and hence base current, a transverse electric field will be set up in the inactive part of the base, which will tend to draw injected carriers away from the vicinities of the emitter and collector junctions. This will increase the probability that these carriers will recombine on the surface. Because a rise in effective lifetime is actually observed in passing to high emitter currents, it seems that this effect is of secondary importance.

For the surface-barrier transistor the effective lifetime measured by the transient method is very much larger than the product CR_1 at all currents. It is also much smaller than the volume lifetime of the base material. Reference to eqn. (5) will show that if the emitter-efficiency term, $D_n n'_p w / D'_p p_n l_n$, is dominant at all currents, CR_1 may bear no relationship to effective lifetime, and in the limit the right-hand side becomes

$_nl_n/2D_nn_p'$. The steady-state results here are therefore com-ible with a dominant emitter-efficiency term. The fall in the ve at high levels is explained by an increasing value for n_p'. e transient results show a rise in effective lifetime at high ection levels, as for alloy-type transistors.

There is an apparent fall in effective lifetime for the types 760 and CK762 transistors in passing from low to inter-diate emitter currents. If this fall is real and not due to an dercorrection for C_E/G, a qualitative explanation for this aviour may be made in terms of saturable traps for minority riers in the base, possibly induced by strains in the crystal ice. These traps would be most effective at low currents and refore give rise to an apparently greater effective lifetime der such conditions. With increase of emitter current the uration of these traps would render them impotent. There is independent experimental evidence to confirm this suggestion p–n–p alloy-type germanium transistors.

n the above discussion it has been assumed that the effective times are predominantly surface-controlled. This would not essarily be true if a marked degradation of the bulk lifetime the base material took place during transistor fabrication, ough an increase of recombination-centre density within the e. This could occur, for instance, if impurity-type recom-ation centres diffused in, or if strains were set up in, the stal. This is no firm evidence on this point, but the work of derhandler and Giacoletto[14] shows that the extent of a possible gradation of bulk lifetime is insufficient to prevent surface atment from influencing the effective lifetime.

(4.2) Effective Diffusion Constant

For transistor types TS2 and CK760 the variation of D_p' is hin the one-dimensional theoretical limit of 2 : 1, although in h case an asymptotic value has not been reached at low rents. By exercising some latitude in selecting the value of to use in correcting the CR_2 plots, it is possible to constrain to a 2 : 1 variation for the other two alloy-type transistors. this is attempted for the surface-carrier transistor, however, corrected curve becomes markedly non-monotonic and has a onounced maximum, as may be seen in Fig. 2(v), so that for this nsistor, by limiting to 2 : 1 the ratio of the values of D_p' at high l low emitter currents, this ratio will be exceeded in the inter-diate current range. The values of C_E that must be used for types EW59 and CK762 transistors for $I_E = 10\,\mu$A, to make value of $CR_2 - C_E/G$ at this current equal to twice the h-level value, are 190 and 52 pF respectively, in contrast to the culated values of 123 and 35 pF. These differences can be plained in terms of a gross departure from the theoretical ationship for C_E in eqn. (29), in so far that the values of V_E uld have to be considerably greater than those measured ectly for the emitter-to-base terminal voltage appropriate for $= 10\,\mu$A.

n some cases, therefore, the one-dimensional theoretical atment of transistor performance can apply to practical nsistors only if an effective lifetime, and also an effective fusion constant that is allowed to vary outside the one-nensional theoretical limits of 2 : 1, replace bulk lifetime and nority-carrier diffusion constant wherever they occur.

A possible explanation of the excessive variation of D_p' may based on the scattering effect of the surface as follows: In ctical alloy-type and surface-barrier transistors there is often onounced curvature of the emitter-base and collector-base erfaces; this, coupled with the effect of the low surface recom-ation velocities that can be obtained, will mean that many riers injected into the base will, in fact, reach the collector y after 3-dimensional diffusion, some of which will be into l out of the surface layers near the emitter and collector. Schrieffer[18] has shown that in surface channels the actual mobility of carriers (and hence their diffusion constant) may be reduced to as little as one-tenth of the bulk value; an average effective diffusion constant for carriers crossing from the emitter to the collector of a transistor can therefore be less than the free bulk value, provided that surface scattering has a pronounced effect on the diffusive motion of a significant number of carriers. It is considered that the effect of the surface in determining an average effective diffusion constant, D_p', would decrease with increasing injection level for the same reason that τ_{pe}' increases.

The explanation for low-to-medium-current behaviour of τ_{pe}' discussed in the previous Section, involving trapping, can also explain the apparently low values of D_p' that are deduced at low emitter currents. The effect of carrier trapping will be to reduce the effective diffusion constant below the free bulk value. With the saturation of the traps at medium and high emitter currents, their effect is nullified.

Although there is no evidence of D_p' reaching an asymptotic value at low currents, in each case such a level was reached at high currents. Such consistency points in favour of accepting that, under these circumstances, the value of D_p' is indeed $2D_p$, as was assumed in Section 2.2.2. In these circumstances the base width, w, may be determined from eqn. (8). Values so obtained are given in Table 1.

Table 1

Base Width, w, calculated from High-Level Values of CR_2

Transistor type	TS2	CK760	CK762	EW59	SB100
w, cm ($\times 10^{-3}$)	6·3	2·8	1·2	4·3	0·51

It may be mentioned here that, because most practical alloy-type transistors have emitter and collector boundaries with the base that are concave inwards, w—which we have so far considered as a well-defined parameter—will not be equal to the minimum emitter-collector spacing. It will be greater than this, because many of the injected carriers in the base will have to traverse a longer path. Because the proportions of injected carriers crossing the base by different routes will depend on the injection level, the weighting of these routes, which determine a mean w, will also depend on the injection level. In other words, w itself may be regarded as an effective parameter rather than an invariant one; it will, in fact, decrease with increasing injection level, because, with the growth of the electric field in the base, a greater proportion of carriers will diffuse between the emitter and the collector by short routes.

However, in interpreting the results it is convenient to consider w fixed in eqn. (6) and to regard the variation of $w^2/2D_p'$ entirely as a variation of D_p'. The expected decrease of w' with increasing injection level discussed above will then be reflected in a somewhat greater variation of D_p' than would be deduced if w' had been allowed to vary as well.

(4.3) Current Gain

(4.3.1) α_{cb0}: Low-Frequency Value between Base and Collector.

For types TS2, CK760 and SB100 the peak value of α_{cb0} is not greater than twice the low-level value. It is known for the surface-barrier transistor that the emitter efficiency is low, even at low emitter currents. This is instrumental in reducing the observed ratio of the maximum value of α_{cb0} to its low-level value to less than 2 : 1, despite the apparently large increase in D_p' in passing from low to high emitter currents for this transistor. The behaviour of the types TS2 and CK760, on the other hand, conforms to the analysis given by Webster.

The EW59 and CK762 transistors show a range much greater than 2 : 1 for variation of α_{cb0}. Reference to eqns. (1) and

(2) will show that, because the major part of the change occurs in a current range within which the emitter efficiency term is small, this change is caused by an increase in the product $D_p'\tau_{pe}'$, or equivalently, since $wv'A_s'/D_p'A \simeq w^2/2D_p'\tau_{pe}'$, by a decrease in $v'A_s'/D_p'$. Reference to the two preceding Sections shows that both D_p' and τ_{pe}' show an overall increase in passing from low to high currents for the EW59 transistor, while for the type CK762 the change is accounted for by variation of D_p' which more than offsets an apparent fall in τ_{pe}'.

(4.3.2) $f_{\alpha d}$: Cut-Off Frequency of α_{cd}.

The internal diffusion current gain, α_{ced}, between the emitter and the collector is given by $-y_{21}/y_{11}^*$, where y_{11}^* is that part of y_{11}, as defined in eqn. (13), that arises from minority-carrier diffusion. $f_{\alpha d}$, the frequency at which $|\alpha_{ced}|$ is reduced to $1/\sqrt{2}$ times its zero-frequency value, is related to D_p' by[5]

$$f_{\alpha d} = \frac{a}{2\pi}\frac{D_p'}{w^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

In terms of the product CR_2 of the present work this may be written as

$$f_{\alpha d} = \frac{a}{4\pi}\left(CR_2 - \frac{C_E}{G}\right)^{-1} \quad . \quad . \quad . \quad . \quad (10)$$

Here a is a numerical coefficient whose precise value is a function of w/l_p', where $l_p' = (D_p'\tau_{pe}')^{1/2}$, but which varies monotonically[5] only from 2·43 for $w/l_p' = 0$ to 2·63 for $w/l_p' = 0{\cdot}44$. A mean value of 2·5 may be taken for a with good approximation for the transistors discussed here. In Table 2 a comparison of the experimental value of $f_{\alpha d}$ and that calculated from eqn. (10) is made for all five transistors at an emitter current of 1 mA. The experimental values of $f_{\alpha d}$ were determined* from measurements of the short-circuit current gain of the transistors using apparatus similar to that described by Coffey.[19] Appropriate corrections were made to the measurements of short-circuit current gain to take account of internal feedback through the series path comprising C_c and r_{b0}—the collector depletion-layer capacitance and the external base resistance respectively—and so yield the internal diffusion-current gain. For the surface-barrier transistor, measurements of the internal diffusion-current gain were not possible beyond 50 Mc/s. It was found, however, that up to this frequency the internal diffusion-current gain followed closely the theoretical form given by

Table 2

CALCULATED AND EXPERIMENTAL VALUES OF $f_{\alpha d}$

Transistor type	$CR_2 - \frac{C_E}{G}$	$f_{\alpha d}$	
		Calculated	Measured
	sec × 10⁻⁹	Mc/s	Mc/s
TS2 ..	275	0·72	0·8
CK760 ..	44	4·5	6·0
CK762 ..	9·0	22	22·2
EW59 ..	110	1·8	1·5
SB100 ..	1·5	130	95

$$-y_{21}/y_{11}^* \simeq \operatorname{sech}\theta$$

where $\theta = w/l_p'(1 + j\omega\tau_p')^{1/2}$. It was therefore possible to determine $f_{\alpha d}$ by an extrapolation based on this expression.

Agreement between the experimental and calculated values of $f_{\alpha d}$ is seen to be reasonably good for the alloy-type transistors. For the surface-barrier transistor the $f_{\alpha d}$ values obtained are considerably higher than those indicated by the manufacturers, which are based on data involving the maximum frequency of oscillation of the transistor. The measure of agreement shown is an indication of the closeness with which the theoretical one-dimensional transistor formula for α_{ced}, modified to include an effective lifetime and diffusion constant deduced from low-frequency measurements, fits experimental high-frequency data.

* HYDE, F. J., and SMITH, R. W. (to be published).

(5) CONCLUSION

It has been demonstrated that the one-dimensional small-signal internal-transistor equations may be adapted to explain the experimentally-observed values of α_{cb0} and $f_{\alpha d}$ for commercial transistors, provided that the bulk lifetime and diffusion constant of minority carriers in the base are replaced by effective values. The value of the effective lifetime to be used in these equations has been shown to be equivalent to that determined by an independent transient method: in general, this effective lifetime increases with injection level. There is evidence that the effective diffusion constant to be used in the modified one-dimensional equations varies outside the one-dimensional theoretical range of 2 : 1.

It is worth noting that, because the assumptions involved in taking a one-dimensional model and in it replacing the key parameters by those governed by 3-dimensional behaviour are so gross, the further simplifying assumptions made by some authors when deducing an equivalent circuit for the practical transistor may be insignificant beside them.

The interpretation of the experimental data has been attempted on the basis of the transistors having uniform current densities across the emitter-to-base interfaces. In a private communication Webster has stated that the following types of imperfection are fairly prevalent in emitter recrystallized regions: (*a*) areas of very thin regrowth; (*b*) areas with no regrowth; (*c*) unwetted spots. These would undoubtedly make any clear-cut analytical deductions impossible, yet they might well account for some of the effects observed.

(6) ACKNOWLEDGMENTS

The author wishes to acknowledge the assistance of his colleagues, Messrs. P. A. Smith, R. W. Smith and R. E. Taylor, in the experimental work.

The work described above was carried out as part of the programme of the Radio Research Board. The paper is published by permission of the Director of Radio Research of the Department of Scientific and Industrial Research.

(7) REFERENCES

(1) SHOCKLEY, W.: 'Electrons and Holes in Semiconductors' (D. Van Nostrand, New York, 1950), p. 318 *et seq.*
(2) SHOCKLEY, W.: 'The Theory of *p–n* Junctions in Semiconductors and *p–n* Junction Transistors', *Bell System Technical Journal*, 1949, **28**, p. 435.
(3) MOORE, A. R., and PANKOVE, J. I.: 'The Effect of Junction Shape and Surface Recombination on Transistor Current Gain, Part I', *Proceedings of the Institute of Radio Engineers*, 1954, **42**, p. 907.
(4) STRIPP, K. F., and MOORE, A. R.: 'The Effects of Junction Shape and Surface Recombination on Transistor Current Gain, Part II', *ibid.*, 1955, **43**, p. 856.
(5) PRITCHARD, R. L.: 'Frequency Variations of Junction Transistor Parameters', *ibid.*, 1954, **42**, p. 786.
(6) GIACOLETTO, L. J.: 'Study of *p–n–p* Alloy Junction Transistors from D.C. through Medium Frequencies', *RCA Review*, 1954, **15**, p. 506.

(7) EVANS, D. M.: 'Measurements on Alloy-Type Germanium Transistors and Their Relation to Theory', *Journal of Electronics*, 1956, **1**, p. 461.
(8) WEBSTER, W. M.: 'On the Variation of Junction Transistor Current-Amplification Factor with Emitter Current', *Proceedings of the Institute of Radio Engineers*, 1954, **42**, p. 914.
(9) RITTNER, E.: 'Extension of the Theory of the Junction Transistor', *Physical Review*, 1954, **94**, p. 1161.
(10) MISAWA, T.: 'Emitter Efficiency of Junction Transistor', *Journal of the Physical Society of Japan*, 1955, **10**, p. 362.
(11) ARMSTRONG, L. D., CARLSON, C. L., and BENTIVEGNA, M.: 'P–N–P Transistors Using High-Emitter-Efficiency Alloy Materials', *RCA Review*, 1956, **17**, p. 37.
(12) LO, A. W., *et al.*: 'Transistor Electronics' (Prentice-Hall Inc., New York, 1955), p. 292.
(13) GOSSICK, B. R.: 'Transient Behaviour of Rectifiers', *Journal of Applied Physics*, 1955, **26**, p. 1356.
(14) LEDERHANDLER, S. R., and GIACOLETTO, L. J.: 'Measurement of Minority Carrier Lifetime and Surface Effects in Junction Diodes', *Proceedings of the Institute of Radio Engineers*, 1955, **43**, p. 477.
(15) HENDERSON, J. C., and TILLMAN, J. R.: 'Minority-Carrier Storage in Semi-Conductor Diodes', *Proceedings I.E.E.*, Paper No. 2293 R, January, 1957 (**104** B, p. 318).
(16) SCHWARZ, R. F., and WALSH, J. F.: 'The Surface Barrier Transistor, Part V. The Properties of Metal to Semiconductor Contacts', *Proceedings of the Institute of Radio Engineers*, 1953, **41**, p. 1715.
(17) MOORE, A. R., and WEBSTER, W. M.: 'The Effective Surface Recombination of a Germanium Surface with a Floating Barrier', *ibid.*, 1955, **43**, p. 427.
(18) SCHRIEFFER, J. R.: 'Effective Carrier Mobilities in Surface Space Charge Layers', *Physical Review*, 1955, **97**, p. 641.
(19) COFFEY, W. N.: 'Measuring Radio Frequency Parameters of Junction Transistors', *Electronics*, 1956, **29**, p. 152.
(20) ZAWELS, J.: 'Physical Theory of New Circuit Representation for Junction Transistors', *Journal of Applied Physics*, 1954, **25**, p. 976.
(21) HORTON, H. V., KOCH, W., RATH, H. L., and SCHULTZ, W.: 'Barrier-Layer Capacitance of the Ge–In Alloyed Rectifier', *Zeitschrift für Physik*, 1954, **138**, p. 336.
(22) MUSS, D. R.: 'Capacitance Measurements on Alloyed Indium-Germanium Junction Diodes', *Journal of Applied Physics*, 1955, **26**, p. 1514.
(23) BOOTHROYD, A. R., and ALMOND, J.: 'A Bridge for Measuring the A.C. Parameters of Junction Transistors', *Proceedings I.E.E.*, Paper No. 1681 M, September, 1954 (**101**, Part III, p. 314).

(8) APPENDICES

(8.1) Conditions for Balance of Bridge

A schematic for the bridge is shown in Fig. 1. In terms of the admittance parameters of the internal transistor (i.e. ignoring ...), which are defined by

$$\left.\begin{aligned} i_e &= y_{11}v_e + y_{12}v_c \\ i_c &= y_{21}v_e + y_{22}v_c \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

the conditions for balance can be shown to be

$$\frac{R_2}{Z} = -\frac{y_{11} + y_{22} + y_{12} + y_{21}}{y_{21} + y_{22}} - \frac{R_2(y_{11}y_{22} - y_{12}y_{21})}{y_{21} + y_{22}} \quad (12)$$

where $1/Z = 1/R_1 + j\omega C$. The derivation of the admittance parameters has been carried out by several authors for a transistor with planar geometry and with unidimensional current flow: the nomenclature of Zawels[20] will be used here with the exception of G, which here refers to an actual transistor of emitter area A. If the collector efficiency is assumed to be unity but the effects of emitter and collector depletion-layer capacitances and junction leakages are included,

$$\left.\begin{aligned} y_{11} &= G\theta \coth\theta + Y_n + Y_{IE} + j\omega C_E \\ y_{12} &= -\frac{G\theta \operatorname{cosech}\theta}{K} \\ y_{21} &= -G\theta \operatorname{cosech}\theta \\ y_{22} &= \frac{G\theta\coth\theta}{K} + Y_{IC} + j\omega C_C \end{aligned}\right\} \quad . \quad . \quad (13)$$

where
$$G = \frac{Ae}{w}\mu_p p_n \varepsilon^{\Lambda V_E}$$

$$\theta = \frac{w}{(D_p'\tau_p')^{1/2}}(1 + j\omega\tau_p')^{1/2}$$

$$\frac{Y_n}{G} = \frac{D_n n_p' w}{D_p p_n l_n}(1 + j\omega\tau_n)^{1/2}$$

$$\frac{1}{K} = \frac{1}{\Lambda(D_p'\tau_p')^{1/2}}\frac{\partial w}{\partial V_C}\operatorname{cosech}\frac{w}{(D_p'\tau_p')^{1/2}} + \frac{\varepsilon^{\Lambda V_C}}{\varepsilon^{\Lambda V_E}}$$

Here Y_{IE} and Y_{IC} are leakage admittances across the emitter and collector junctions, μ_p is the mobility of holes, τ_n is the lifetime of electrons in the emitter, C_E and C_C are the emitter and collector depletion-layer capacitances respectively, and $\Lambda = e/kT$. In practice, for good commercial transistors at low frequencies the admittances y_{12} and y_{22} can be ignored, as can the emitter leakage admittance Y_{IE}, so that the condition for balance is closely approximated by

$$\frac{R_2}{R_1}(1 + j\omega CR_1) = -\frac{y_{11} + y_{21}}{y_{21}} = \frac{1}{\alpha_{cb}} \quad . \quad . \quad (14)$$

if R_2 is chosen to be small.

The effect of surface recombination on this balance condition will be considered by two methods. The first embodies surface recombination in an effective lifetime, τ_{pe}', which replaces τ_p' in the above equations, while in the second the effect of surface recombination is retained as a separate factor. This second treatment was initiated by Webster[8] for d.c. or zero-frequency conditions: in effect, a separate input admittance of the form $Y_{SR} = Gwv'A_s'/D_p'A$ is added to the one-dimensional admittance y_{11}. Lo *et al.*[12] have suggested an extension to a.c. conditions which consists in multiplying Y_{SR} by a factor $(1 + j\omega\tau_p')^{1/2}$.

(8.1.1) Use of Effective Lifetime.

For this case

$$\frac{R_2}{R_1}(1 + j\omega CR_1) = \cosh\theta - 1 + \frac{(Y_n + j\omega C_E)\sinh\theta}{G\theta} \quad (15)$$

The low-frequency expansion of this for $w/[D_p'\tau_{pe}']^{1/2} \ll 1$ is given by

$$\frac{R_2}{R_1}(1 + j\omega CR_1) = \left(\frac{1}{2}\frac{w^2}{D_p'\tau_{pe}'} + \frac{D_n n_p' w}{D_p' p_n l_n}\right) + j\omega\left(\frac{\tau_{pe}'}{2}\frac{w^2}{D_p'\tau_{pe}'} + \frac{C_E}{G}\right) \quad . \quad . \quad . \quad . \quad (16)$$

Therefore

$$CR_1\left(1 - \frac{1}{CR_2}\frac{C_E}{G}\right) = \frac{w^2}{2D_p'}\left(\frac{w^2}{2D_p'\tau_{pe}'} + \frac{D_n n_p' w}{D_p' p_n l_n}\right)^{-1} \quad . \quad (17)$$

$$CR_2 - \frac{C_E}{G} = \frac{w^2}{2D_p'} \quad . \quad . \quad . \quad . \quad . \quad . \quad (18)$$

$$\frac{R_1}{R_2} = \left(\frac{1}{2}\frac{w^2}{D_p'\tau_{pe}'} + \frac{D_p'n_p'w}{D_p'p_nl_n}\right)^{-1} \quad . \quad . \quad . \quad . \quad (19)$$

Note that these expressions differ from those given by Evans,[7] in that the effects of injection efficiency and emitter capacitance are included.

G is proportional to emitter current, whereas C_E varies much more slowly with emitter current; in consequence, the term C_E/G is important for currents of the order of a few hundred microamperes and below for practical transistors. For larger currents this term is negligible in practice, although theoretically C_E will rise very rapidly as the emitter-to-base p.d. approaches 200 mV. For practical high currents eqns. (17) and (18) become

$$CR_1 = \frac{w^2}{2D_p'}\left(\frac{1}{2}\frac{w^2}{D_p'\tau_{pe}'} + \frac{D_nn_p'w}{D_p'p_nl_n}\right)^{-1} \quad . \quad . \quad . \quad (20)$$

$$CR_2 = \frac{w^2}{2D_p'} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (21)$$

If the level of injection is not too high,

$$\frac{D_nn_p'w}{D_p'p_nl_n} \ll \frac{w^2}{2D_p'\tau_{pe}'}$$

so that $CR_1 \simeq \tau_{pe}'$.

(8.1.2) Use of Separate Surface-Recombination Admittance.

The form of the admittance will be taken as that suggested by Lo *et al.*,[12] namely $Y_{SR}(\omega) = \frac{Gwv'A_s'}{D_p'v'A}(1 + j\omega\tau_p')^{1/2}$. From eqns. (13) and (14),

$$\frac{R_2}{R_1}(1 + j\omega CR_1) = \cosh\theta - 1 + \frac{(Y_n + Y_{SR} + j\omega C_E)\sinh\theta}{G\theta} \quad (22)$$

The low-frequency approximation of this, for $w/(D_p'\tau_p')^{1/2} \ll 1$ is

$$\frac{R_2}{R_1}(1 + j\omega CR_1) = \left(\frac{1}{2}\frac{w^2}{D_p'\tau_p'} + \frac{D_nn_p'w}{D_p'p_nl_n} + \frac{wv'A_s'}{D_p'A}\right) + j\omega\tau_p'\left(\frac{1}{2}\frac{w^2}{D_p'\tau_p'} + \frac{1}{2}\frac{wv'A_s'}{D_p'A} + \frac{C_E}{G\tau_p'}\right) \quad . \quad (23)$$

so that

$$CR_1 = \tau_p'\left(\frac{1}{2}\frac{w^2}{D_p'\tau_p'} + \frac{1}{2}\frac{wv'A_s'}{D_p'A} + \frac{C_E}{G\tau_p'}\right)\left(\frac{1}{2}\frac{w^2}{D_p'\tau_p'} + \frac{D_nn_p'w}{D_p'p_nl_n} + \frac{wv'A_s'}{D_p'A}\right)^{-1} \quad (24)$$

$$CR_2 = \tau_p'\left(\frac{1}{2}\frac{w^2}{D_p'\tau_p'} + \frac{1}{2}\frac{wv'A_s'}{D_p'A} + \frac{C_E}{G\tau_p'}\right) \quad . \quad (25)$$

$$\frac{R_1}{R_2} = \left(\frac{1}{2}\frac{w^2}{D_p'\tau_p'} + \frac{D_nn_p'w}{D_p'p_nl_n} + \frac{wv'A_s'}{D_p'A}\right)^{-1} \quad . \quad . \quad (26)$$

As before, these results may be simplified for currents at which the term C_E/G is unimportant, provided that

$$\frac{D_nn_p'w}{D_p'p_nl_n} \ll \frac{w^2}{2D_p'\tau_p'}$$

when eqns. (24) and (25) become

$$CR_1 = \tau_p'\left(\frac{1}{2}\frac{w^2}{D_p'\tau_p'} + \frac{1}{2}\frac{wv'A_s'}{D_p'A}\right)\left(\frac{1}{2}\frac{w^2}{D_p'\tau_p'} + \frac{wv'A_s'}{D_p'A}\right)^{-1} \quad (27)$$

$$CR_2 = \frac{w^2}{2D_p'}\left(1 + \frac{vA_s'\tau_p'}{Aw}\right) \quad . \quad . \quad . \quad . \quad (28)$$

(8.1.3) Discussion of Above Conditions.

The fraction R_1/R_2 of eqns. (19) and (26) is simply the zero-frequency value of α_{cb}, namely α_{cb0}. There is a marked difference, however, in the two expressions for the product of CR_1 of eqns. (17) and (24) which arises because of the particular form of the frequency-dependence of $Y_{SR}(\omega)$ suggested by Lo *et al.* The difference may be clearly seen if the emitter efficiency and emitter capacitance terms can be ignored, when from eqn. (20) $CR_1 \simeq \tau_{pe}'$, while from eqn. (27), even though v' is infinity, the product CR_1 cannot theoretically be less than $\tau_p'/2$. Evans's results[7] for a type TA153 transistor under suitable conditions show that CR_1 is of the order of 10 microsec, which is very much smaller than half of the bulk lifetime of the base material of this transistor; Webster quoted 500 microsec for τ_p. In view of this it seems that the a.c. extension to Webster's theory proposed by Lo *et al.* is inappropriate, unless a marked degradation of bulk lifetime had, in fact, taken place during transistor fabrication.

(8.2) Correction for Emitter Depletion-Layer Capacitance

For step junctions such as occur in alloy-type and surface-barrier transistors the expression for emitter depletion-layer capacitance is

$$C_E = A\left[\frac{\epsilon\epsilon_0 e\,(N_d - N_a)}{2(V_0 - V_E)}\right]^{1/2} \quad . \quad . \quad . \quad (29)$$

in M.K.S. units. Here ϵ is the relative permittivity of the base material (16 for germanium) $\epsilon_0 = 8\cdot86 \times 10^{-12}$ farad/metre, $e = 1\cdot6 \times 10^{-19}$ coulomb, $N_d - N_a$ is the net donor density in the base and V_E is the applied emitter junction voltage. V_0 is a constant under reverse bias conditions, and for the order of base resistivities involved in the present transistors is about 0·2 volt[21, 22] at room temperature. Measurements of C_E with large reverse bias were made for the alloy-type transistors, using an admittance bridge operating at 100 kc/s. The value of V_0 above was confirmed. Values of C_{E0}, the emitter depletion-layer capacitance for $V_E = 0$ are given, with an estimated accuracy of 5%, in Table 3.

Table 3

VALUES OF C_{E0} FOR ALLOY-TYPE TRANSISTORS

Transistor type	TS2	CK760	CK762	EW59
C_{E0}, pF ..	22	35	31	120

For the surface-barrier transistor, a value for C_{E0} of about 4·5 pF was estimated from manufacturers' data. Stray capacitance prevented a direct measurement for this transistor.

In Section 8.1 the correction term C_E/G arises. Theoretically, $G \simeq eI_E/kT$, provided that I_E is considerably greater than the reverse saturation current, I_S, of the emitter junction. The applicability of this relationship was verified, down to the smallest emitter currents used, from low-frequency measurements of the hybrid parameters, using a bridge of the type described by Boothroyd and Almond.[23] The actual value of C_E for each emitter current was determined from eqn. (29), which was assumed to be valid for forward (positive) values of V_E using, instead of V_E, the measured d.c. value of the emitter-to-base terminal voltage appropriate for the particular value of I_E. For small values of I_E this voltage should approximate to V_E, because the voltage drop across the external base resistance r_{b0} is very small.

For the types CK762 and EW59 transistors the measured values of the emitter-to-base voltage were somewhat smaller than the values of V_E expected from the theoretical relationship

$$I_E = I_S\,[\exp(eV_E/kT) - 1] \quad . \quad . \quad . \quad . \quad (30)$$

in which measured values of I_S were used.

1.317.38.029.64

The Institution of Electrical Engineers
Paper No. 2450 R
Jan. 1958

THE HALL EFFECT AND ITS APPLICATION TO POWER MEASUREMENT AT 10 Gc/s

By Professor H. E. M. BARLOW, Ph.D., B.Sc.(Eng.), Member, and S. KATAOKA.

(*The paper was first received* 14*th June, and in revised form* 31*st August*, 1957.)

SUMMARY

The paper describes experiments on the measurement of microwave wer at 10 Gc/s employing the Hall effect produced in single crystals n- and p-type germanium when erected on the axis of a hollow etal rectangular waveguide carrying the power. So far as is known, is is the first recorded observation of the Hall effect at this frequency, d it has been shown possible to apply the effect, with the help of a itable phase-adjustment device, to the design of a wattmeter. For is purpose the Hall output from a crystal was calibrated against the wer measured independently, and the relationship was found to be actically linear.

From these power measurements and consideration of the conditions operation of the crystal, the Hall coefficients for the *n*- and *p*-types germanium used were deduced and shown to be of the same order magnitude as the d.c. values.

An investigation of the impedance of the crystal circuit led to the nclusion that the contribution to the Hall effect from the displace-ent current in the crystal was very small compared with that of the nduction current.

The success of this application depends largely upon the development suitable crystal units, and towards that end new techniques were troduced.

LIST OF PRINCIPAL SYMBOLS

ϵ_r = Relative permittivity.
σ = Conductivity.
τ_n, τ_p = Mean free time for electrons and holes, respectively.
i = Instantaneous current in crystal unit; direction parallel to applied electric field.
$= \mu_0 h$ = Instantaneous flux density in crystal unit.
v = Instantaneous Hall e.m.f.
e = Instantaneous electric field.
S = Instantaneous Poynting vector, $= e \times h$.
I, B, E = Complex r.m.s. values of i, b and e respectively, appropriate to sinusoidal signal.
$\bar{v}, \bar{S}$ = Time average values of v and S.
$\mathscr{R}$ = Hall coefficient.
t = Thickness of crystal.
Z = Shunt wave impedance of crystal.
Z_0 = Wave impedance of H_{01}-mode.
β = Phase coefficient of H_{01}-mode $[=2\pi/\lambda_g]$.
l = Distance of short-circuit from crystal.
ρ = Reflection coefficient.
P = Total power flowing along the waveguide.

(1) INTRODUCTION

In previous papers[1-4] the application of the Hall effect in mi-conductors to power measurement has been discussed and xperimental work on this development up to 4 Gc/s has been scribed. So far as the authors are aware, no observations ave previously been made on this effect at frequencies as high 10 Gc/s, although it is known[5,6] that the conductivity and permittivity of germanium at these frequencies differ very little from the corresponding d.c. values.

Benedict and Shockley[7] measured the permittivity of high-purity germanium at 24 Gc/s and found $\epsilon_r = 18$–19 at room temperature. Experiments carried out by Klinger[8] at the same frequency show that the conductivity of pure germanium is then about half the d.c. value, but the relative permittivity remains practically unchanged at about 15·5.

According to the theory developed by Kronig,[9] the microwave conductivity of a semi-conductor is given by

$$\sigma(\omega) = \frac{\sigma_{d.c.}}{1 + (\omega\tau)^2}$$

with $\begin{cases} \tau = \tau_n \text{ for } n\text{-type} \\ \tau = \tau_p \text{ for } p\text{-type} \end{cases}$

for extrinsic conditions and by:

$$\sigma(\omega) = \frac{1}{2}\sigma_{d.c.}\left[\frac{1}{1 + (\omega\tau_p)^2} + \frac{1}{1 + (\omega\tau_n)^2}\right]$$

for intrinsic conditions, where ω is the angular frequency of the microwave.

Since the mean free time is the order of 10^{-12} sec for any high-purity germanium, it is to be expected that the microwave conductivity at 10 Gc/s would be only slightly smaller than the d.c. value. At this frequency the displacement current in the semi-conductor becomes at least as large as the conduction component, but there has so far been no evidence that it contributes substantially to the Hall effect. It is important, however, to bear in mind that only that part of the displacement current which arises from the presence of the material medium, and is therefore additional to the free-space current, can subscribe to the Hall effect.[4]

The experiments described in the paper were designed in the first place to establish the existence of the Hall effect at 10 Gc/s, and, if that were successful, to apply the effect to measurements of the corresponding power. In the course of the work it was hoped to get some information about the contribution of displacement current to the Hall e.m.f.

(2) THEORY OF THE METHOD

If a small slab of semi-conductor is placed at the centre of a rectangular waveguide supporting the dominant mode, as shown in Fig. 1, so that the electric field, e, produces in the crystal an instantaneous current, i, a Hall e.m.f. will be set up along the x-axis, owing to the interaction of this current with the transverse magnetic field of the wave in the guide. The current i is shown as a vector quantity, its direction being the same as that of the applied electric field e.

With uniform distribution of current and magnetic field over the small cross-section of the slab, the instantaneous Hall e.m.f. is given by

$$v = \left(\frac{\mathscr{R}}{t}\right)|i \times b|$$

Written contributions on papers published without being read at meetings are vited for consideration with a view to publication.
Professor Barlow is Pender Professor of Electrical Engineering at University College, M. Kataoka, who was formerly at University College, London, is now at the iversity of Tokyo, Japan.

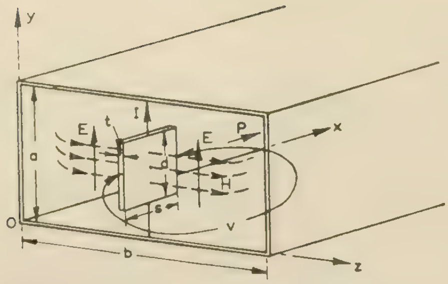


Fig. 1.—Principle of wattmeter.

The time average of the e.m.f. is therefore

$$\bar{v} = \mathrm{Re}\left[\left(\frac{\mathscr{R}}{t}\right) I \times B^*\right] \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where B^* is the complex conjugate of B.

Since the Poynting vector of power flux density is $S = E \times H$ and its time average is $\bar{S} = \mathrm{Re}\,[E \times H^*]$, H^* being the complex conjugate of H, $\bar{v}$ will be directly proportional to $\bar{S}_x$ when the values of I and B are suitably related to those of E and H respectively, and when the directions of the vector quantities are those shown in Fig. 1. In particular, it is important to ensure proper adjustment of the phase of the current through the crystal with respect to the microwave magnetic field, and this is conveniently established when the time average of the Hall e.m.f. is zero with the waveguide terminated in a short-circuit. The device should then operate as a wattmeter for power passing along the waveguide, this power being equal to $\bar{S}_x$ integrated over the guide cross-section.

To provide the necessary conditions, a small slab of germanuim crystal is erected at a point on the axis of a waveguide, as shown in Fig. 2, and wires attached in the E-plane to the crystal are

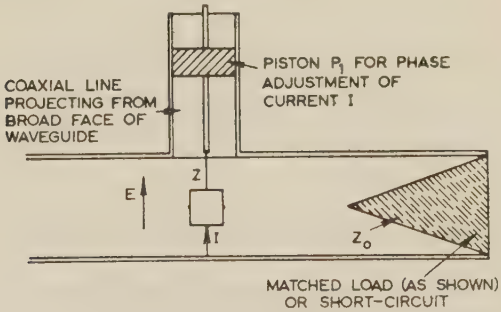


Fig. 2.—Crystal circuit.

joined, on one side directly to the wall of the guide and on the other to the centre conductor of a coaxial line having an adjustable piston at its end for the purpose of phasing the current through the crystal.

Suppose that the effect of the crystal with its associated coaxial line is equivalent to a shunt wave impedance Z across the waveguide at the point at which it is erected. Then the total impedance at this same point, looking from the generator to the load when the waveguide is terminated at a distance l in a short-circuit, is

$$Z_s = \frac{jZZ_0 \tan \beta l}{Z + jZ_0 \tan \beta l} \quad . \quad . \quad . \quad . \quad . \quad (2)$$

and the corresponding reflection coefficient is

$$\rho_s = \frac{-Z + j(Z - Z_0)\tan \beta l}{Z + j(Z + Z_0)\tan \beta l} \quad . \quad . \quad . \quad (3)$$

When the waveguide has a matched termination the impedance Z_L and reflection coefficient ρ_L looking towards the load from the crystal will be

$$Z_L = \frac{ZZ_0}{Z + Z_0} \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

$$\rho_L = \frac{-Z_0}{2Z + Z_0} \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

This reflection coefficient for a matched load differs from zero owing to the presence of the crystal circuit.

When an electric field E_y^+ is incident on the crystal circuit, a current, I, is induced in it and this produces a secondary electric field, E_y^-, related to the total current I through the crystal by the equation[10]

$$E_y^- = -\sqrt{\left(\frac{\mu_0}{\epsilon_0}\right)} \frac{\lambda_g}{\lambda} \frac{I}{b} \sin\left(\frac{\pi z}{b}\right) \varepsilon^{\mp j\beta x} \quad . \quad . \quad . \quad (6)$$

This assumes that the crystal and the wires attached to it are small, so that they can be regarded as a straight filament stretched across the waveguide and causing only a slight disturbance of the field in it.

Now, in respect of the dominant mode, E_y^- is related to E_y^+ by the reflection coefficient ρ_L with a matched load beyond the crystal (so that ρ_L arises from the crystal only) and

$$E_y^- = \rho_L E_y^+ \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

Thus, since the crystal is on the axis of the waveguide and at $x = 0$,

$$I = -\rho_L E_y^+ \sqrt{\left(\frac{\epsilon_0}{\mu_0}\right)} \frac{\lambda}{\lambda_g} b \quad . \quad . \quad . \quad . \quad . \quad (8)$$

With ρ_L small, so that the field in the waveguide remains relatively undisturbed, the total power along the guide is

$$P = E_y^+ H_z^+ \left(\frac{ab}{2}\right) \quad . \quad . \quad . \quad . \quad . \quad (9)$$

and

$$\frac{E_y^+}{H_z^+} = Z_0 = \frac{\lambda_g}{\lambda}\sqrt{\frac{\mu_0}{\epsilon_0}} \quad . \quad . \quad . \quad . \quad (10)$$

giving

$$P = \frac{1}{2}(E_y^+)^2 \frac{\lambda}{\lambda_g}\sqrt{\left(\frac{\epsilon_0}{\mu_0}\right)} ab \quad . \quad . \quad . \quad . \quad (11)$$

Since the contribution of the displacement current in the crystal to the Hall effect is unknown, we assume Hall coefficients $\mathscr{R}_c$ for the effect arising from the conduction component and $\mathscr{R}_d$ for the corresponding effect arising from the displacement component.

The ratio of these Hall coefficients is

$$r = \frac{\mathscr{R}_d}{\mathscr{R}_c} \quad . \quad . \quad . \quad . \quad . \quad . \quad (12)$$

Then, remembering that it is only the displacement current through the material medium that is effective in supplementing the conduction current, the equation for the time average of the Hall e.m.f. will be modified to become

$$\bar{v} = \mathrm{Re}\left[\frac{\mathscr{R}_c}{t} \frac{\sigma + j\omega r(\epsilon - \epsilon_0)}{\sigma + j\omega\epsilon} I B_z^*\right] \quad . \quad . \quad . \quad (13)$$

If E_y^+ is purely real, then, for a matched load,

$$B_z^* = \mu_0 H_z^+ = E_y^+ \sqrt{(\mu_0 \epsilon_0)} \frac{\lambda}{\lambda_g} \quad . \quad . \quad . \quad (14)$$

Combining eqns. (11), (13), (14) and (8), the time average of ae Hall e.m.f. for a matched load is

$$\bar{v}_L = -\frac{\mathscr{R}_c}{t}(E_y^+)^2 b\epsilon_0\left(\frac{\lambda}{\lambda_g}\right)^2 \mathrm{Re}\left[\rho_L\frac{\sigma + j\omega r(\epsilon - \epsilon_0)}{\sigma + j\omega\epsilon}\right]$$

r $$\bar{v}_L = -\frac{2\mathscr{R}_c}{at}\frac{\lambda}{\lambda_g}\sqrt{(\epsilon_0\mu_0)}P\,\mathrm{Re}\left[\rho_L\frac{\sigma + j\omega r(\epsilon - \epsilon_0)}{\sigma + j\omega\epsilon}\right] \qquad (15)$$

nd is directly proportional to the power.

Thus, if the reflection coefficient, ρ_L, for a matched load, the Hall e.m.f., $\bar{v}_L$, and the power, P, at a given frequency are ieasured, eqn. (15) enables the Hall coefficient, $\mathscr{R}_c$ (for conducion current) to be calculated.

If I_c is the conduction component of the current through the rystal whose dimensions are s, t and d (see Fig. 1), the power issipated in it is

$$P_c = |I_c|^2\left(\frac{d}{\sigma st}\right)$$

here $$I_c = \left(\frac{\sigma}{\sigma + j\omega\epsilon}\right)I$$

Vhen the waveguide is terminated in a matched load, eqns. (8) nd (11) can be used to find the ratio of the power consumed in ae crystal to that in the load, i.e.

$$\frac{P_c}{P} = \frac{2|\rho_L|^2 bd}{\sigma ast}\sqrt{\left(\frac{\epsilon_0}{\mu_0}\right)}\frac{\lambda}{\lambda_g}\left(\frac{\sigma^2}{\sigma^2 + \omega^2\epsilon^2}\right) \quad . \quad . \quad (16)$$

Now suppose that the waveguide is short-circuited by a piston t a distance l from the crystal not precisely an integral number f quarter wavelengths, so that the shunt impedance of the crystal high compared with the impedance presented by the terminaon of the guide. Measuring positive values of l from the shortrcuit, we get standing waves of electric and magnetic field, long the axis of the guide, represented by

$$\left.\begin{aligned} E_S &= j2E_y^+ \sin\beta l \\ H_S &= 2H_z^+ \cos\beta l \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (17)$$

ith magnetic flux density

$$B_S = \mu_0 H_S = 2E_y^+\sqrt{(\mu_0\epsilon_0)}\frac{\lambda}{\lambda_g}\cos\beta l \quad . \quad . \quad (18)$$

he corresponding current induced in the crystal circuit is

$$I_S = -j2\rho_L E_y^+\sqrt{\left(\frac{\epsilon_0}{\mu_0}\right)}\frac{\lambda}{\lambda_g}b\sin\beta l \quad . \quad . \quad (19)$$

ince $B_S^* = B_S$, the time average of the Hall e.m.f. when the aveguide is short-circuited is given by

$$\bar{v}_S = -\frac{4b\mathscr{R}_c\epsilon_0}{t}(E_y^+)^2\left(\frac{\lambda}{\lambda_g}\right)^2\sin\beta l\cos\beta l$$

$$\mathrm{Re}\left[j\rho_L\frac{\sigma + j\omega r(\epsilon - \epsilon_0)}{\sigma + j\omega\epsilon}\right] \quad . \quad (20)$$

ut $\bar{v}_S = 0$ for proper adjustment of the phasing piston, so that

$$\mathrm{Re}\left[j\rho_L\frac{\sigma + j\omega r(\epsilon - \epsilon_0)}{\sigma + j\omega\epsilon}\right] = 0$$

r $$\mathrm{Im}\left[\rho_L\frac{\sigma + j\omega r(\epsilon - \epsilon_0)}{\sigma + j\omega\epsilon}\right] = 0 \quad . \quad . \quad . \quad (21)$$

nd $\left[\rho_L\frac{\sigma + j\omega r(\epsilon - \epsilon_0)}{\sigma + j\omega\epsilon}\right]$ must be wholly real, which simplifies ne interpretation of eqn. (15).

ρ_L can be measured, and if its value is

$$\rho_L = \rho' + j\rho'' \quad . \quad . \quad . \quad . \quad . \quad . \quad (22)$$

then $$r = \frac{\rho' - \rho''\left(\frac{\sigma}{\omega\epsilon}\right)}{\rho'\left(\frac{\epsilon - \epsilon_0}{\epsilon}\right) + \rho''\frac{\omega(\epsilon - \epsilon_0)}{\sigma}} \quad . \quad . \quad . \quad (23)$$

and the corresponding value of $\mathscr{R}_d$ can be calculated. When $l \simeq n(\lambda_g/4)$ and the impedance presented at the crystal by the termination of the waveguide is large compared with Z, then even under these conditions of a short-circuited waveguide a residual Hall output will appear, given by

$$\bar{v}_S' = -\frac{b\mathscr{R}_c}{t}(E_y^+)^2\epsilon_0\left(\frac{\lambda}{\lambda_g}\right)^2$$

$$\mathrm{Re}\left[\rho_L\frac{\sigma + j\omega r(\epsilon - \epsilon_0)}{\sigma + j\omega\epsilon}(1 - \rho_0^*)(1 + \rho_0)\right] \quad (24)$$

where $\rho_0 = \dfrac{Z - Z_0}{Z + Z_0}$

It is therefore desirable, when adjusting the piston associated with the coaxial line in the crystal circuit, to give correct phasing of the current I, to avoid the condition $l \simeq n(\lambda_g/4)$, but in any case if Z is large the error is small.

(3) EXPERIMENTAL ARRANGEMENTS AND THEIR APPLICATION TO THE DESIGN OF A HALL-EFFECT WATTMETER

To examine the behaviour of a small germanium crystal when mounted in a waveguide, experiments were made using a standard waveguide bench. A 9 360 Mc/s microwave generated by a c.w. magnetron was passed in turn through a variable water-tube attenuator, a wavemeter, a standing-wave indicator and the transmission-type Hall-effect wattmeter under investigation, being finally absorbed by a matched load or reflected by a short-circuiting piston. The wattmeter unit itself is shown in Fig. 3,

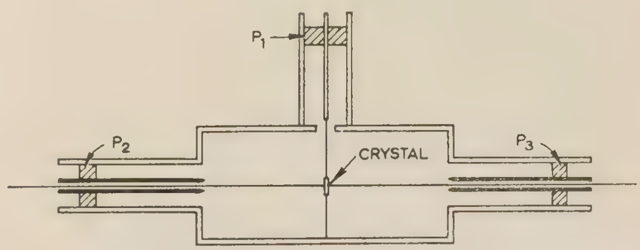


Fig. 3.—Transmission-type Hall-effect wattmeter unit.

where P_1 is the coaxial-line piston for the adjustment of the phase of the current in the crystal, and P_2 and P_3 are pistons associated with coaxial-line Hall output leads so as to give the highest possible h.f. impedance to these leads and at the same time to eliminate residual rectifier action.

One of the current leads to the crystal was attached to a point on the wall of the waveguide immediately opposite it and the other to the centre conductor of the coaxial line with piston P_1. The Hall leads, on the other hand, were d.c. insulated from the centre conductors of the coaxial lines with pistons P_2 and P_3. To establish the correct positions for these three coaxial-line pistons the following procedure was adopted.

First, each piston was placed at a distance of about $3/4\lambda$ from the opening of the corresponding coaxial line into the waveguide,

which itself was terminated in another short-circuiting piston, thereby setting up a pure standing wave with no power passing along the waveguide.

The pistons P_2 and P_3 were then reset to reduce, so far as possible by this means, the output from the Hall leads, and this was followed by an adjustment of the phasing of the current through the crystal by the piston P_1 to bring the output at the Hall terminals to a minimum for all positions of the short-circuiting piston at the end of the waveguide.

In principle, the condition to be established is represented by eqn. (24), but in practice there are spurious effects which make it difficult to achieve this precisely. Thus there is almost inevitably a superimposed residual rectifier output at the Hall leads, but this can be distinguished because, unlike the true Hall effect, it does not change sign with reversal of the direction of the power in the waveguide. Another effect, to which Stephenson has called attention* as being the cause of possible disturbance when the guide supports a large standing wave, is unbalance of the thermo-electric e.m.f. at the Hall contacts on the crystal. If there is a gradient of electric field stationary in space across the dimension s of the crystal then there is a tendency for one of the Hall contacts to be heated more than the other, giving a differential thermo-electric e.m.f. Semi-conductor-to-metal contacts exhibit a very large thermo-electric effect, which is about 400 μV/deg C with pure germanium,[11] and consequently even a small temperature gradient in the crystal may be significant. Moreover, when the section of waveguide in which the crystal is erected is turned end for end, keeping the position of the crystal in the standing-wave pattern the same, the temperature gradient in the crystal will be reversed and therefore any

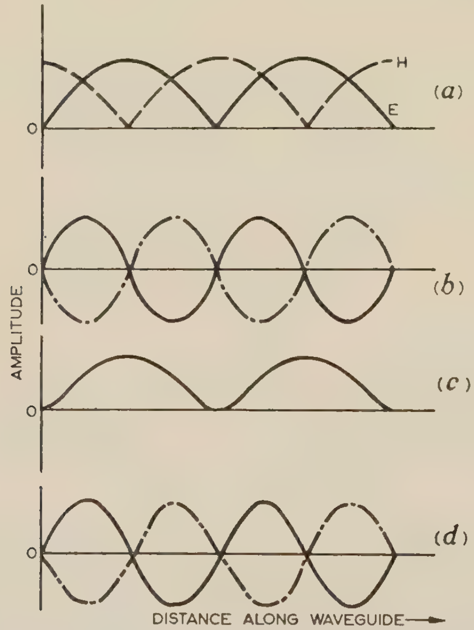


Fig. 4.—Output patterns for waveguide short-circuited at termination.

—— Normal arrangement of crystal unit in waveguide.
—·— Crystal unit turned end for end in waveguide.
(a) Electric- and magnetic-field components of standing wave.
(b) Output from Hall e.m.f.
(c) Output from residual rectifier action.
(d) Output from differential thermal e.m.f.

differential thermo-electric e.m.f. will also be reversed. In these circumstances, the true Hall effect similarly changes sign. Thus, the three components of output at the Hall terminals for a pure standing wave in the waveguide are as shown in Fig. 4.

* Further details will be published later.

Thermal e.m.f. may be distinguished from true Hall e.m.f. by following the variation of output from the crystal that accompanies changes in the phasing of the current through it when the piston P_1 is moved. The thermal e.m.f. depends only on the gradient of the square of the electric-field component of the standing wave in the guide, and this is comparatively little influenced by the position of P_1, while the Hall e.m.f. changes its sign with change of phase of the conduction component of the crystal current in relation to the magnetic field. A metal-foil electric screen covering each end of the crystal to which the Hall leads are attached practically eliminates any thermal e.m.f. The rectifier output, however, is a more difficult problem, but residual effects from this source can usually be reduced to a low level by a final readjustment of the pistons P_2 and P_3 after P_1 has been set to give the correct phasing.

The output voltage at the Hall terminals was measured by a photocell galvanometer-amplifier having a sensitivity of 10·3 mm/μV with an input impedance of 100 kilohms. When the waveguide was terminated by a matched load, the instrument was calibrated against the power absorbed by a differential thermal wattmeter.[12]

The success of this application depends very largely on the preparation of the semi-conductor element. It is obviously desirable to make the crystal as small as possible and of high impedance, so as to avoid disturbing the field in the waveguide more than necessary.

Measurements were made on a number of n and p-type germanium crystals of various resistivities, but the best results were obtained with a very small n-type germanium single crystal and a capacitance-coupled gold-diffused p-type germanium single crystal. The following particulars refer to these.

(4) PREPARATION OF CRYSTALS

(4.1) *n*-Type Germanium Single Crystal

A small slice cut from a single-crystal ingot of n-type germanium having a resistivity of 11 ohm-cm was ground with Carborundum powder in water and then etched in HNO_3–HF solution to give a final size of 0·88 mm long, 0·5 mm wide and 0·125 mm thick. Four 41 s.w.g. tinned-copper wires were soldered with the help of an organic flux to the centre of each edge of the crystal. To do this, the four wires were stretched side by side supported by two pieces of spring metal and the crystal was inserted transversely between the wires so as to separate them. Sufficient current was then passed through the wires to melt the surface tinning at the points of contact with the crystal. Using this procedure, each soldered area was less than 0·02 mm^2. After attachment of the leads to the crystal, it was carefully cleaned in hot water and in alcohol. The final measured resistances between opposite faces of the crystal were 1 950 and 1 450 ohms respectively.

(4.2) *p*-Type Capacitance-Coupled Gold-Diffused Germanium Crystal

The need to provide a high-impedance crystal circuit has already been pointed out: this reduces not only the power absorbed by the crystal, but also the rectifier action. At the same time, if a high-resistivity crystal is used, the displacement current becomes a larger proportion of the whole, so that, to meet the required phasing condition, it is necessary to introduce into the external crystal circuit a correspondingly large capacitive reactance. This is not easily achieved with the piston P_1. By making capacitance connections to the crystal for the current leads, not only are the required phasing conditions more easily satisfied, but rectifier action is avoided at these contacts and a screen is provided to eliminate any thermal e.m.f. by maintaining

uniform electric-field distribution over the ends of the crystal, even when a large standing-wave exists in the waveguide.

Based on these principles, a capacitance-coupled crystal was prepared, a special technique being employed to reduce to negligible proportions any residual rectifier action at the Hall contacts. A slice of *p*-type germanium having a resistivity of 15 ohm-cm was cut to dimensions of 1·47 mm × 0·625 mm × 0·20 mm. After careful surface treatment by etching, a small amount of gold-gallium alloy was diffused into the crystal over a narrow region at both ends, to provide non-rectifying contacts for the Hall leads, as shown in Fig. 5.

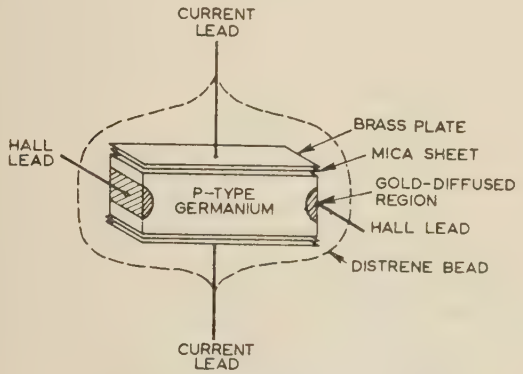


Fig. 5.—Gold-diffused capacitance-coupled crystal unit.

The diffusion process was carried out in a high vacuum at about 850° C for 10–15 min, a small pulse of current being passed through the junction to promote the diffusion process and ensure good electrical non-rectifying contacts. On each of these gold-diffused areas a 41 s.w.g. tinned-copper wire was soldered as a Hall lead.

For the current contacts, a small brass plate of dimensions 1·5 mm × 0·25 mm × 0·025 mm was attached, with a piece of thin mica as a separator, to each of the longer edges of the crystal; 41 s.w.g. wires were then soldered to these condenser plates to form the current leads, the capacitance of each connection being estimated as about 1 pF. After completing the assembly, which was carried out with the help of a micromanipulator, the whole unit was covered with Distrene varnish and dried.

(5) EXPERIMENTAL RESULTS

(5.1) *n*-Type Germanium Crystal

As a result of a careful analysis of the output pattern from the Hall terminals for a pure standing-wave in the guide, no thermal e.m.f. was detected, but there was often some residual rectifier output which could be eliminated at any particular power level by readjusting the pistons P_2 and P_3.

After the correct setting of P_1, P_2 and P_3 had been established the Hall output was calibrated against the power absorbed by a matched load. As shown in Fig. 6, the result was a linear

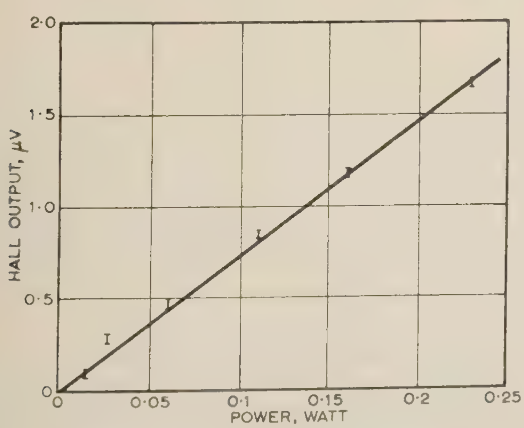


Fig. 6.—Calibration curve for Hall-effect wattmeter using *n*-type germanium crystal.

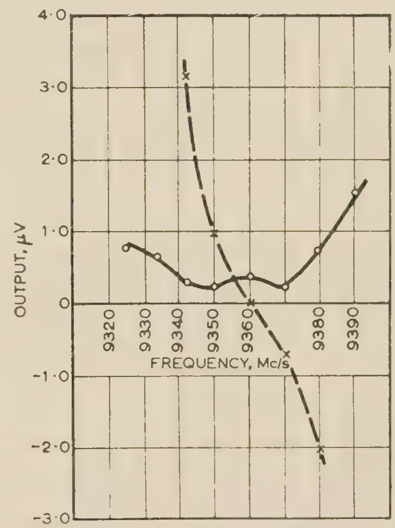


Fig. 7.—Frequency characteristic for *n*-type-germanium Hall-effect wattmeter operating on matched load.

—— Hall output.
---- Output from rectifier action.

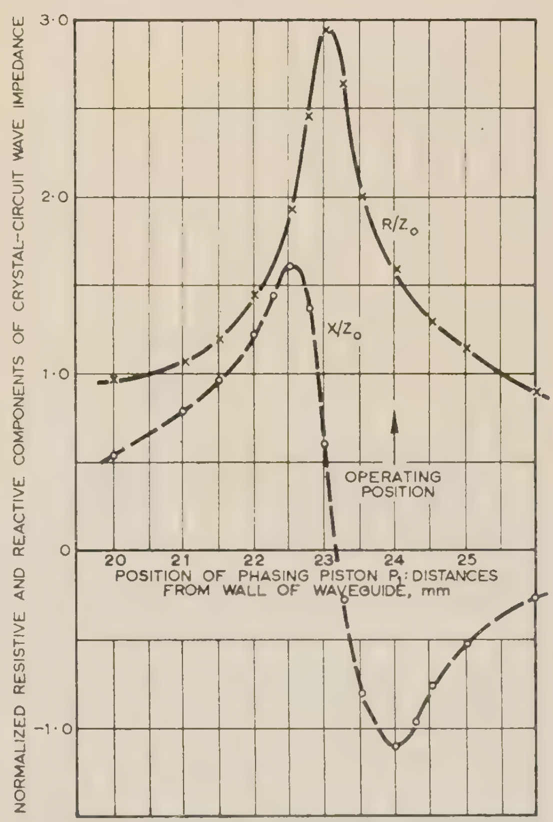


Fig. 8.—Resistive and reactive components of wave impedance at the crystal, with waveguide open-circuited beyond that point.

—— R/Z_0 } for *n*-type germanium.
---- X/Z_0 }

relationship, and the proportionality constant was determined as $7 \cdot 5\,\mu V$/watt. At high power levels the Hall output tends to rise more slowly, and this can be attributed to heating of the crystal. Employing, as it does, three tuning pistons, the device was extremely frequency-sensitive. Fig. 7 shows the Hall and the rectifier outputs over a range of frequencies when all pistons were set for one particular value.

As a means of elucidating the behaviour of the crystal in the waveguide, measurements were made with a standing-wave indicator to obtain the impedance at the point at which the crystal was mounted, with an open-circuit beyond it. The resistive and reactive components of this impedance are shown in Fig. 8 for different settings of the phasing piston. The arrow on the curve indicates the normal operating position, which is clearly very capacitive.

(5.2) Capacitance-Coupled *p*-Type Germanium Crystal

Fig. 9 shows the output at the Hall terminals for a waveguide short-circuited at the end with a piston in different positions.

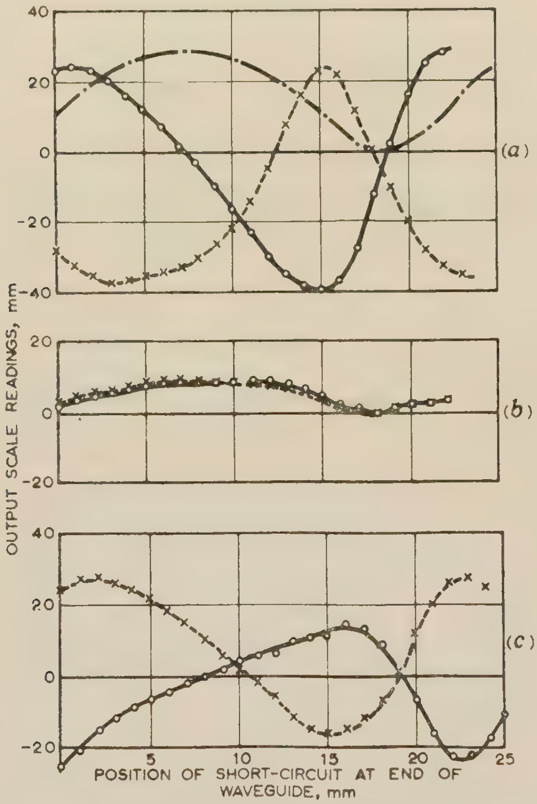


Fig. 9—Patterns of output from capacitance-coupled germanium crystal with waveguide short-circuited.

——— Normal arrangement of crystal.
– – – – Crystal unit turned end for end in waveguide.
—·— Electric field distribution.

(*a*) Conduction current in crystal lagging applied magnetic field; P_1 is 23 mm from waveguide wall.
(*b*) Conduction current in crystal in phase with applied magnetic field; P_1 is $23 \cdot 5$ mm from waveguide wall.
(*c*) Conduction current in crystal leading applied magnetic field; P_1 is 24 mm from waveguide wall.

Curves are given for three different settings, (*a*), (*b*) and (*c*), of the phasing piston P_1 and in each case for the section of waveguide in which the crystal was mounted connected in circuit first one way and then turned end for end. The correct position of the phasing piston was found to be at about $23 \cdot 5$ mm, which is significantly shorter than the corresponding position for the *n*-type crystal. It will be observed that the output patterns were inverted, as would be expected for the P_1 piston settings larger and smaller than $23 \cdot 5$ mm. No thermal e.m.f. was observed, and the small amount of rectifier action could be made negligible by slight readjustment of the Hall pistons P_2 and P_3. The general performance of this crystal was an improvement on the *n*-type. Fig. 10 shows the calibration of the instrument giving

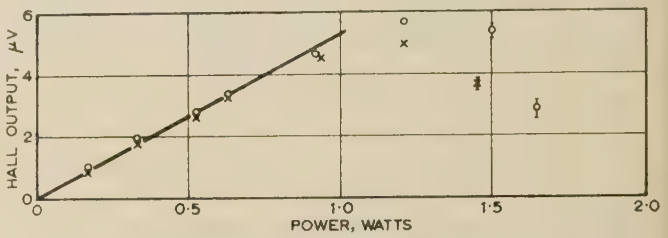


Fig. 10.—Calibration curve for Hall-effect wattmeter using capacitance-coupled *p*-type-germanium crystal.

● ● ● Normal arrangement of crystal.
× × × Crystal unit turned end for end in waveguide.

Hall output against power. Below 1 watt a reasonably linear relationship was obtained, but at higher power levels the Hall output began to decrease and the device became unstable as the power approached 2 watts. The measured wave impedance at the crystal with the waveguide open-circuited beyond it, was $Z/Z_0 = 2 \cdot 21 - j2 \cdot 64$. Turning the wattmeter unit end-for-end did not affect the output at the Hall terminals seriously, but there was a slight difference between measurements of Z/Z_0 in the two cases. This was probably due partly to residual rectifier action at the Hall contacts and partly to asymmetry in the mounting of the crystal.

(6) CALCULATIONS FROM THE THEORY AND ANALYSIS OF RESULTS

(6.1) *n*-Type Germanium Crystal

The skin depth for *n*-type germanium of resistivity 11 ohm-cm and relative permittivity 16 at 9 360 Mc/s is $2 \cdot 53$ mm, and the ratio of conduction current to displacement current is $\sigma/\omega\epsilon = 1 \cdot 1$.

The majority-carrier (electron) density at room temperature is $1 \cdot 58 \times 10^{20}$ per cubic metre (assuming an electron mobility of 3 600 cm^2/volt-sec) and the corresponding minority-carrier (hole) density is $3 \cdot 73 \times 10^{18}$ per cubic metre. Therefore the contribution of minority carriers to the Hall coefficient is in this case negligible, and the theoretical Hall coefficient becomes

$$\mathscr{R} = \frac{3\pi}{8}\left(\frac{1}{Nq}\right) = 4 \cdot 6 \times 10^{-2}\,\mathrm{m^3/coulomb}$$

From the experimental data given in Fig. 8 we find the normalized wave impedance of the crystal circuit to be $Z/Z_0 = 1 \cdot 56 - j1 \cdot 09$ and its reflection coefficient for a matched waveguide termination to be $\rho_L = -(0 \cdot 19 + j0 \cdot 1)$. Putting these values into eqn. (23) gives $r = 0 \cdot 35$.

Now we can proceed to calculate from eqn. (15) the value of $\mathscr{R}_c$ and from eqn. (16) the value of P_c/P using the following figures appropriate to this crystal:

$$\bar{v}_L/P = 7 \cdot 5\,\mu\mathrm{V/watt}$$
$$a = 0 \cdot 01\,\mathrm{m}$$
$$b = 0 \cdot 023\,\mathrm{m}$$
$$t = 0 \cdot 000\,125\,\mathrm{m}$$
$$d = 0 \cdot 000\,5\,\mathrm{m}$$

$$s = 0{\cdot}000\,88\,\mathrm{m}$$
$$\lambda_g = 0{\cdot}044\,8\,\mathrm{m}$$
$$\lambda = 0{\cdot}032\,\mathrm{m}$$
$$\frac{1}{\sqrt{(\mu_0\epsilon_0)}} = 3 \times 10^8\,\mathrm{m/sec}$$
$$\rho_L = -(0{\cdot}19 + j0{\cdot}1),\ \text{with } r = 0{\cdot}35.$$

Thus we find:

$$\mathscr{R}_c = 1{\cdot}20 \times 10^{-2}\,\mathrm{m^3/coulomb\ and\ } P_c/P = 10\%$$

It will be observed that this value of the Hall coefficient at nearly 10 Gc/s is of the same order of magnitude as the theoretically calculated d.c. value. Moreover, the contribution from displacement current in the crystal to the Hall effect is small compared with that of the conduction current. But in interpreting these results it is important to bear in mind that the theory is only approximate, and the conclusions drawn from these calculations may be more qualitative than quantitative. Errors may be introduced, because

(*a*) Temperature rise of the crystal during operation may produce more electrons and consequently decrease the Hall coefficient.

(*b*) The theory assumed that the crystal and the wires attached to it could be represented by a very thin filament. In fact, although the crystal was exceedingly small, it was probably large enough to give a significant reflection at the surface with the incident wave and this tends to increase ρ_L whilst decreasing the apparent $\mathscr{R}_c$ value.

(*c*) The crystal circuit was assumed throughout to have a comparatively high impedance, and the effect of the Hall leads on that impedance was neglected. Measurements showed that Z was not really high enough fully to justify the assumption made.

(6.2) Capacitance-Coupled *p*-Type Germanium Crystal

The skin depth for *p*-type germanium of resistivity 15 ohm-cm and relative permittivity 16 at 9 360 Mc/s is 3·40 mm, and $\sigma/\omega\epsilon = 0{\cdot}81$.

Taking the hole mobility as 1 700 cm²/volt-sec, the majority-carrier (hole) density at room temperature as $2{\cdot}46 \times 10^{20}$ per cubic metre and the minority-carrier density ($2{\cdot}54 \times 10^{18}$) as negligible, we find the theoretical d.c. Hall coefficient to be $\cdot0 \times 10^{-2}\,\mathrm{m^3/coulomb}$. Since in this case $Z/Z_0 = 2{\cdot}21 - j2{\cdot}64$, we get $\rho_L = -0{\cdot}0945 - j0{\cdot}0925$ and $r = 0{\cdot}10$, so that the contribution of the displacement current to the Hall effect is relatively very small.

Using these values in eqns. (15) and (16) together with

$$\bar{v}_L/P = 5{\cdot}4\,\mu\mathrm{V/watt}$$
$$a = 0{\cdot}01\,\mathrm{m}$$
$$t = 0{\cdot}000\,2\,\mathrm{m}$$
$$d = 0{\cdot}000\,625\,\mathrm{m}$$

and
$$s = 0{\cdot}001\,47\,\mathrm{m}$$

we find the Hall coefficient to be $\mathscr{R}_c = 2{\cdot}7 \times 10^{-2}\,\mathrm{m^3/coulomb}$ and $P_c/P = 2\%$. This figure for $\mathscr{R}_c$ shows a closer correspondence to the theoretical d.c. value than did the equivalent figure for *n*-type germanium.

The improved performance and better agreement with theory in the present instance are considered to be due to the higher impedance of the capacitance-coupled type of crystal unit and the lower dissipation in it.

(7) CONCLUSIONS

The experimental observations described demonstrate quite clearly the existence of a Hall effect in both *n*- and *p*-type single crystals of germanium at 10 Gc/s. Hall coefficients deduced from these experiments are of the same order of magnitude as the corresponding theoretical d.c. values.

The contribution of the displacement component of the current in the crystal to the Hall effect was found to be small compared with that of the conduction current.

This transmission type of Hall-effect wattmeter gives a reasonably good performance for small power levels (below about 1 watt), but it is seriously disturbed by temperature rise of the crystal when higher powers are used. This drawback can be overcome by using a directional coupler to feed the instrument when a large power is to be measured. The use of semi-conductors other than germanium may also lead to an improvement. The frequency sensitivity of the instrument can probably be diminished by including a resonant iris to reduce the dependence of the crystal impedance on frequency, but this has not yet been investigated.

(8) ACKNOWLEDGMENTS

The authors acknowledge with gratitude the most helpful discussions of this work with Professor A. L. Cullen, Dr. J. Brown, Mr. H. G. Effemey and Dr. L. M. Stephenson. They also wish to thank the British Council and the Trustees of the R. W. Paul Fund for financial support.

(9) REFERENCES

(1) Barlow, H. E. M.: 'The Application of the Hall Effect in a Semi-Conductor to the Measurement of Power in an Electromagnetic Field', *Proceedings I.E.E.*, Paper No. 1654 M, June, 1954 (**102 B**, p. 179).

(2) Barlow, H. E. M.: 'The Design of Semi-Conductor Wattmeters for Power-Frequency and Audio-Frequency Applications', *ibid.*, Paper No. 1778 M, November, 1954 (**102 B**, p. 186).

(3) Barlow, H. E. M., and Stephenson, L. M.: 'The Hall Effect and its Application to Power Measurement at Microwave Frequencies', *ibid.*, Paper No. 1913 R, January, 1956 (**103 B**, p. 110).

(4) Barlow, H. E. M.: 'Hall Effect and its Counterpart, Radiation Pressure, in Microwave Power Measurement', *ibid.*, Monograph No. 191 R, August, 1956 (**104 C**, p. 35).

(5) Donovan, B., and March, N. H.: 'High-Frequency Conductivity in Semi-Conductors', *Proceedings of the Physical Society*, 1956, **69**, p. 528.

(6) D'Altroy, F. A., and Fan, H. I.: 'Effect of Neutral Impurity on the Microwave Conductivity and Dielectric Constant of Germanium at Low Temperature', *Physical Review*, 1956, **103**, p. 1671.

(7) Benedict, T. B., and Shockley, W.: 'Microwave Observations of the Collision Frequency of Electrons in Germanium', *ibid.*, 1953, **89**, p. 1152.

(8) Klinger, Y.: 'The Conductivity of Germanium at $2{\cdot}4 \times 10^{10}$ c/s', *ibid.*, 1953, **92**, p. 509.

(9) Kronig, R. de L.: 'The Quantum Theory of Dispersion in Metallic Conductors', *Proceedings of the Royal Society*, A, 1931, **133**, p. 255.

(10) Lewin, L.: 'Advanced Theory of Waveguides' (Iliffe, London, 1951).

(11) Dunlap, W. C.: 'An Introduction to Semi Conductors' (Wiley, New York, 1957), p. 257.

(12) Gordon-Smith, A. C.: Private communication.

(10) APPENDIX

When the impedance of the part of the crystal circuit represented by the coaxial line with its associated wires is very high, the current through the crystal will be supplemented by the direct action of the electric field in the waveguide and will be

partly collected as a displacement current from the surrounding air dielectric.

In these circumstances, for a matched load, the supplementary crystal current is approximately given by

$$I' = ts(\sigma + j\omega\epsilon)KE_y^+$$

where K is the ratio between electric field within the crystal and the field outside it.

For a small spherical crystal of germanium placed in a static and uniform electric field,

$$K \simeq \frac{3}{\epsilon_r + 2} = \frac{1}{6}$$

and for a similar crystal of infinite size having an electromagnetic wave incident upon it,

$$K \simeq \frac{2}{\sqrt{(\epsilon_r)} + 1} = \frac{2}{5}$$

In the present experiment the crystal was extremely thin, so that the former case will be more nearly applicable. The total current in the crystal is then

$$I = -\rho_L E_y^+ \sqrt{\left(\frac{\epsilon_0}{\mu_0}\right)} \frac{\lambda}{\lambda_g} b + ts(\sigma + j\omega\epsilon)KE_y^+ \quad . \quad (25)$$

and the corresponding Hall e.m.f. is

$$\bar{v} = \frac{\mathscr{R}_c}{t} E_y^+ \operatorname{Re}\left\{\left[-\rho_L \sqrt{\left(\frac{\epsilon_0}{\mu_0}\right)} \frac{\lambda}{\lambda_g} b + ts(\sigma + j\omega\epsilon)K\right] \frac{\sigma + j\omega r(\epsilon - \epsilon_0)}{\sigma + j\omega\epsilon} B_z^*\right\}$$

so that, for a matched load,

$$\bar{v}_L = \frac{\mathscr{R}_c}{t} \sqrt{(\mu_0\epsilon_0)} \frac{\lambda}{\lambda_g} (E_y^+)^2 \operatorname{Re}\left\{\left[-\rho_L \sqrt{\left(\frac{\epsilon_0}{\mu_0}\right)} \frac{\lambda}{\lambda_g} b + tsK(\sigma + j\omega\epsilon)\right] \frac{\sigma + j\omega r(\epsilon - \epsilon_0)}{\sigma + j\omega\epsilon}\right\}$$

For the n-type germanium crystal used, we have

$$-\rho_L \sqrt{\left(\frac{\epsilon_0}{\mu_0}\right)} \frac{\lambda}{\lambda_g} b = (8 \cdot 3 + j4 \cdot 4) \times 10^{-6}$$

taking $\rho_L = -(0 \cdot 19 + j0 \cdot 1)$ and $b = 0 \cdot 023$ m

and $$tsK(\sigma + j\omega\epsilon) = (3 \cdot 3 + j3 \cdot 0) \times 10^{-7}$$

taking $K = \frac{1}{3}$.

Hence, under the conditions of these experiments the supplementary current collected from the air surrounding the crystal is not really significant and can be neglected.

1.317.784

The Institution of Electrical Engineers
Paper No. 2461 M
Jan. 1958

A PRECISION THERMO-ELECTRIC WATTMETER FOR POWER AND AUDIO FREQUENCIES

By J. J. HILL, B.Sc., Associate Member.

(*The paper was first received* 6*th July, and in revised form* 24*th September*, 1957.)

SUMMARY

The output e.m.f. given by a conventional thermal convertor is not oportional to the square of the heater current, and wattmeters using o such convertors cannot therefore be used for the precise measure-ent of power. The causes of non-compliance with a square law are amined, and a means of compensation is given to provide a con-rtor system in which the output e.m.f. is within 0·1% of the cal-lated square-law value.

A wattmeter has been constructed in which negative-feedback ampli-rs are used to supply the currents to two compensated convertors. e input signal voltages to the amplifiers are obtained from the ternal circuit by means of a precision wide-frequency-band voltage nsformer, a 4-terminal non-inductive current shunt and a high-sistance voltage divider.

The insertion loss in the main current circuit is 0·1 volt at rated rrent, and the voltage-divider resistance is either 1 000 or 5 000 ohms r volt. The overall instrument error for all conditions of load and wer factor is 0·1% over the frequency range 200 c/s to 10 kc/s, and 3% for the range 50 c/s to 30 kc/s.

LIST OF PRINCIPAL SYMBOLS

Properties of the heater and thermo-couple wires of a thermal nvertor.

ρ_0 = Resistivity of heater at temperature T_0, ohm-cm.
α = Temperature coefficient of resistivity.
k_0 = Thermal conductivity of heater at T_0, (watts/cm²)/(deg C/cm).
β = Temperature coefficient of thermal conductivity.
p = Perimeter of heater wire, cm.
a = Cross-section of heater, cm².
l = One-half of heater length, cm.
ε = Emissivity coefficient of surface of heater.
σ = Stefan's constant, watts/cm² deg K⁴.
T_0 = Ambient temperature, deg K.
R_c, R_h = Resistance of thermocouple and heater wires respectively at T_0, ohms.
I = Heater current, amp.
θ_M = Mid-point temperature rise of heater, deg C.
V = Output e.m.f. of thermocouple, volts.

(1) INTRODUCTION

When the measurement of power has to be carried out over a equency band covering some decades, or when the fundamental equency of the supply voltage has a large harmonic content, e conventional form of dynamometer wattmeter is not capable precision accuracy, particularly when the circuit power factor low. Electronic amplifiers have been used to supply one or th dynamometer coils in order to reduce the errors at the ghest frequencies, and various instruments designed for parti-lar types of operation have been described.[1,2,3,4,5] In one these[5] an accuracy of 0·2% of full-scale deflection at unity wer factor, and 0·6% at zero power factor, was achieved over the whole of the frequency range 50 c/s–20 kc/s. The upper limit of 20 kc/s was set mainly by the performance of the indicating dynamometer instrument. In order to extend this limit, or to reduce the errors, it would be necessary to design a dynamometer having a resonant frequency of at least 300 kc/s. This might be possible, but only at the cost of unduly reducing the instrument torque. In addition, the design of constant-current amplifiers to operate at higher frequencies with the variable inductive load of a dynamometer movement would introduce further difficulties.

The use of thermal convertors, with their well-known excellent frequency characteristics, as measuring devices in place of dynamometers, therefore has a great attraction. A new instrument has recently been constructed at the N.P.L. in which the usual defects and errors associated with thermal convertors have been virtually eliminated. The accuracy achieved for all conditions of load and power factor is approximately 0·1% over the frequency range 200 c/s–10 kc/s, and 0·3% for the range 50 c/s–30 kc/s.

(2) THE THERMAL CONVERTOR AS A POWER-MEASURING DEVICE

Thermal convertors, in which the e.m.f. developed at the terminals of a thermo-junction in thermal contact with an electric heater gives a measure of the current in the heater, have long been used for the measurement of power. In a thermal wattmeter two convertors are used, one of which is heated by the algebraic sum of two derived currents and the other by their difference. If one of the derived currents is proportional to the circuit current and the other proportional to the voltage, the sum of the outputs of the two thermocouples in series opposition is proportional to power, provided that the two convertors are matched and that the output of each is proportional to the square of the heater current.

In practice, however, thermal convertors do not obey this simple square-law relationship and consequently when they are used in a wattmeter errors up to several per cent may result.

(3) THE CAUSES OF ERROR IN THERMAL CONVERTORS

The characteristics of more than 30 commercially made vacuum-enclosed insulated thermal convertors were measured, and the results of a representative sample are reproduced in Table 1. In order to make more evident the effect of non-compliance with a square law upon the output e.m.f., the final column in the Table gives the quantity S:

$$S = \frac{\text{Output at rated current}}{\text{Output at low current}} \times \left(\frac{\text{Low current}}{\text{Rated current}}\right)^2$$

The value of the low current was that which was necessary to give an output of about 1 mV.

For all convertors tested, having heater materials of any one of the four common resistance alloys quoted in Table 1, the value of S was less than 0·98, the average being 0·95. In order to account for the non-linear relation between the output and the

Written contributions on papers published without being read at meetings are invited for consideration with a view to publication.
The paper is an official communication from the National Physical Laboratory.

Table 1

Heater conditions: Material	Heater conditions: Current rating, I	Heater conditions: Rated voltage drop	Change in heater resistance from $I/3$ to I	Rated output e.m.f.	Mean temperature-coefficient of output	S
	mA	V	%	mV	% per + 1° C ambient*	
Nichrome	1·25	1·85	+2·0	8·4	−0·50	0·820
Nichrome	2·5	1·0	−2·0	7·6	−0·35	0·914
Nichrome	10	0·24	+1·2	6·7	−0·15	0·962
Nichrome	25	0·26	+1·1	9·1	−0·10	0·943
Constantan	50	0·15	−0·3	6·9	−0·10	0·965
Constantan	120	0·14	−0·5	7·0	—	0·970
Constantan	500	0·15	—	7·0	−0·10	0·950
Constantan	1 000	0·17	—	6·5	—	0·965
Manganin	35	0·20	+1·4	8·2	−0·15	0·935
Platino-iridium	25	0·25	+1·0	7·0	−0·10	0·970

* In most cases the temperature coefficient was found to be dependent upon the heater current.

square of the current, it is necessary to examine the factors which may be responsible and hence attempt to assess their relative importance. The principal causes of error are (*a*) the temperature coefficients of resistivity and thermal conductivity of the heater material, (*b*) the radiation losses from the heater surface, and (*c*) the non-linearity of the e.m.f./temperature relationship of the attached thermocouple. An approximate analysis has been carried out (see Appendix) to evaluate the mid-point temperature rise of the heater of a thermal convertor. The results show that for heater materials of high-resistivity alloys this temperature rise, θ_M, can be given by

$$\theta_M = \left(\frac{R_c}{R_c + R_h}\right)\frac{I^2\rho_0 l^2}{2a^2k_0}\left[1 + \frac{5}{6}\alpha\frac{I^2\rho_0 l^2}{2a^2k_0}\left(1 - \frac{5}{6}\beta\frac{I^2\rho_0 l^2}{2a^2k_0}\right)\right.$$
$$- \frac{5}{6}\beta\frac{I^2\rho_0 l^2}{2a^2k_0}\left(1 - \frac{5}{6}\beta\frac{I^2\rho_0 l^2}{2a^2k^0}\right) - \frac{5}{3}\,\frac{p\varepsilon\sigma T_0^3 l^2}{ak_0}\left(1 - \frac{5}{6}\beta\frac{I^2\rho_0 l^2}{2a^2k_0}\right)$$
$$\left. - \frac{11}{5}\,\frac{p\varepsilon\sigma T_0^2 l^2}{ak_0}\,\frac{I^2\rho_0 l^2}{2a^2k_0}\left(1 - \frac{5}{6}\beta\frac{I^2\rho_0 l^2}{2a^2k_0}\right)^2\right] \quad . \quad . \quad . \quad . \quad (1)$$

This equation shows that, as the heater current is increased from a very small value, the mid-point temperature rise differs by an increasing amount from the value to be expected from the simple square-law relationship. The significances of the various terms are as follows:

(i) The factor $R_c/R_c + R_h$ gives a measure of the cooling produced by the attached thermocouple. This may be considerable in low-current convertors, where the heater wire is usually thinner than the thermocouple wires.

(ii) The error due to the term which is dependent upon the temperature coefficient of resistivity, α, rarely exceeds 2% for the materials commonly used in the heaters of thermal convertors.

(iii) The error due to the term which is dependent upon the temperature coefficient of thermal conductivity, β, is usually the largest in all convertors except those of the smallest current rating and may be as great as 20%.

(iv) The remaining terms together give a measure of the errors due to radiation losses, which may amount to 10%. It will be noticed that these terms, which are proportional to T_0^3 and $T_0^2 I^2$, provide a factor giving a convertor a temperature coefficient of output e.m.f. which is current dependent.

The net effect of these errors often results in a departure from the simple square-law relationship by as much as 20% at rated current.

The temperature rise is, however, of interest only in so far as it is the cause of the output e.m.f. which is produced at the terminals of the attached thermocouple. The law relating the output e.m.f., V, and the temperature of a thermocouple is generally given by

$$V = [\theta_1 - \theta_0]\left[X + 273Y + \frac{Y}{2}(\theta_1 + \theta_0)\right] \quad . \quad . \quad . \quad (2)$$

where X and Y are coefficients dependent upon the thermo-electric power of the two metals forming the junction, and θ_1 and θ_0 the cold and hot junction temperatures respectively in degrees centigrade. Putting $\theta_M = \theta_1 - \theta_0$, and considering θ_0 to remain constant, eqn. (2) may be rewritten

$$V = A\theta_M(1 + \gamma\theta_M) \quad . \quad . \quad . \quad . \quad (3)$$

For the alloys commonly used in thermo-junctions, the coefficient A usually amounts to 40 or 50 μV/deg C and γ has values ranging from -2×10^{-4} to 11×10^{-4}, both dependent upon the materials used. It will be noted that the output e.m.f. is dependent upon the ambient temperature θ_0 and therefore this provides a second reason for the temperature coefficient of a thermal convertor. In a typical thermal convertor the output e.m.f., V, at rated current is usually in the range 6–8 mV, which corresponds to a mid-point temperature rise of about 150° C. The effect of γ changes the output e.m.f. for this temperature rise by amounts up to 15% from that which would be obtained from a linear relationship. An error additional to that due to the behaviour of the heater has therefore to be allowed in considering the overall characteristic of a thermal convertor.

In order to obtain V in terms of I^2, the value of θ_M from eqn. (1) is substituted into eqn. (3). The expression is rather unwieldy, and it is not proposed to reproduce it here in full. It has been evaluated for several convertors, and the calculated outputs were found to agree with the measured values to an accuracy of a few per cent. Better accuracy is not possible owing to the uncertainty in determining the values of the parameters involved. The main use of the expression is for assessing the approximate magnitudes of the various error terms as the dimensions, electrical and thermal properties of the heater are changed. Sufficient accuracy for this purpose can be obtained by using a simpler expression of the form

$$V = A\frac{R_c}{R_c + R_h}\,\frac{I^2\rho_0 l^2}{2a^2k_0}[1 + (L\alpha - M\beta - NT_0^2 + P\gamma)I^2] \quad (4)$$

Examination of various types of convertor with resistance-alloy heaters showed that the combined errors due to $M\beta$ and NT_0^2 outweighed those due to $L\alpha + P\gamma$.

The characteristics of about 30 convertors rated from $1\frac{1}{4}$ to

)00 mA were determined experimentally, and it was found that e actual curves could be matched to an accuracy of about $0\cdot1\%$ empirical laws of the form

$$V \simeq KI^2(1 - \Delta I^2) \quad . \quad . \quad . \quad . \quad (4a)$$

nere K and Δ are constants. At rated current the values of ^{2} ranged from $0\cdot02$ to greater than $0\cdot1$. In well-constructed nvertors of medium rating, ΔI^2 is of the order of $0\cdot05$ at rated rrent. If two such convertors were used in a thermal watteter which was adjusted to be correct at rated current and ltage and unity power factor, the scale errors at $\frac{3}{4}$, $\frac{1}{2}$ and $\frac{1}{4}$ ed current would be 1%, 2% and $2\frac{1}{2}\%$ respectively. The ange in error at rated current and voltage and $0\cdot5$ power ctor would be 2%, and the change in error at half load due to a % variation in voltage at constant power would be $0\cdot5\%$. In der to obtain an accuracy of $0\cdot1\%$, ΔI^2 must be smaller than 0025.

COMPENSATION FOR THE RESPONSE ERRORS OF A THERMAL CONVERTOR

It can be seen from eqn. (4) that V would be proportional to if $(L\alpha + P\gamma) = (M\beta + NT_0^2)$. It is doubtful whether a sistance material could be found having such electrical and ermal properties that a heater made of it and allied with the rrect type of thermocouple would give a thermal convertor sely obeying the law $V \propto I^2$. There are, however, materials nich have a very large positive temperature coefficient of resisity. When such materials are used as heaters in thermal nvertors the effect of this large temperature coefficient outighs all the other effects and results in a law of the form

$$V_1 \simeq K_1I^2(1 + \Delta_1I^2) \quad . \quad . \quad . \quad . \quad . \quad (5)$$

ow, if the heaters of two convertors, one responding to eqn. (4a) d the other to eqn. (5), were connected in series, the combined tput $(V + V_1)$ would be proportional to the square of the rrent, provided that $K\Delta = K_1\Delta_1$. Such an identity would be ficult to achieve without careful construction, and selection om a number of convertors would be necessary. But when e heater of one convertor is shunted so that it carries say $1/n$ the current of the other, the combined output is proportional the square of the current, provided that $n = 4\sqrt{(K_1\Delta_1/K\Delta)}$. practical means of compensation for the response errors is ereby indicated. The value of n can be determined approxiately from a knowledge of the constants involved, or more ecisely by experimental trial and error.

The materials selected as being suitable for the heater of a mpensating' convertor were platinum and nickel, and contors rated at 25 and 35 mA were commercially constructed. hen these were tested, ΔI^2 at rated current was found to be the order of $0\cdot3$. Various 'compensated' convertors were en assembled having current ratings of 10 and 25 mA. Each nsisted of either a 10 mA or a 25 mA alloy heater convertor spectively in series with a shunted pure-metal heater convertor. e values of heater resistance and output e.m.f. at various ater currents having been accurately measured for each dividual convertor, it was possible to calculate the value of unting resistance required for any particular combination.

In all, six such combinations were tested, and in each case it s found that the total output e.m.f. of the compensated conr system departed from the calculated square-law value no more than $0\cdot5\%$ at rated current, i.e. the net value of ΔI^2 not exceed $0\cdot005$.

Successive trial adjustments to the values of the shunting ances were made, and usually two were sufficient to reduce errors to about $\pm0\cdot1\%$.

In no case was it found possible to obtain a shunting resistance that would reduce the errors to zero for all values of current, correct compensation at rated current resulting in slight overcompensation at the lower currents. This is due partly to the effects that are dependent upon the fourth and sixth powers of the current and partly to the change in shunting effect consequent upon the increase in resistance of the pure-metal heater with current. In order to reduce the latter effect the ratio of current in the shunt to current in the heater should be as small as practicable, and because the relative increase of output with current of a convertor with a pure-metal heater is greater than the relative decrease of one with an alloy heater, it is desirable that the rating of the pure-metal heater should be higher than that of the alloy heater. Experiments indicate that a ratio of about 2 : 1 is satisfactory.

It might be expected from eqn. (1) and (2) that in a fully compensated convertor the temperature coefficient of output e.m.f. would be reduced to zero. Although this ideal condition was not realized, the coefficient was reduced to about one-fifth of that of the uncompensated alloy convertors. Detailed figures of the performance of two compensated convertors are given in a later Section.

(5) STABILITY OF THERMAL CONVERTORS

It has been shown that it is possible to produce a compensated thermal convertor system closely obeying the square-law relationship and having a very small temperature coefficient of output e.m.f. These advantages would be of little use in the precise measurement of power unless they were accompanied by a reasonably high degree of stability.

(5.1) Short-Period Stability of Alloy-Heater Convertors

The short-period stability of thermal convertors is not usually regarded as being of a high order. Day-to-day changes up to $0\cdot2\%$ and, to a small extent, minute to minute changes, have been observed. The author believes that the stability is really reasonably high and that the changes observed are due to variations in ambient temperature coupled with a high temperature coefficient of output e.m.f.

A 10 mA convertor which had shown an apparent secular stability no better than $0\cdot2\%$ was therefore immersed in a well-stirred oil bath thermostatically controlled to $\pm0\cdot05°$ C. The convertor had a Nichrome heater, and its temperature coefficient of output e.m.f. was $-0\cdot15\%$ per 1° C rise. After the convertor had been immersed for 1 hour its output e.m.f. was measured at rated current at frequent intervals over a period of one week. No change exceeding $0\cdot5\,\mu$V, equivalent to better than $0\cdot01\%$, was observed at any time. It might therefore be expected that a convertor with a very small temperature coefficient would be stable to better than $0\cdot1\%$ when used under normal operating conditions in air.

(5.2) Long-Period Stability of Alloy-Heater Convertors

It is well known that large momentary overloads may change the calibration of a thermal convertor by amounts up to several per cent. If, however, such conditions are avoided a reasonable degree of stability is obtained. The progressive change in the calibration of about a dozen alloy-heater convertors which have been used in precision transfer measurements has not exceeded $0\cdot5\%$ in a period of 10 years.

(5.3) Stability of Compensated Convertors

A compensated convertor was assembled in which a convertor having a Nichrome heater was connected in series with

one having a platinum heater. The temperature coefficient of output e.m.f. of the combination was less than 0·03% per deg C. It was tested at intervals in air at 20 ± 2° C over a period of more than one year and no change in calibration exceeding 0·1% was observed on any occasion.

(6) SPECIFICATION AND DESCRIPTION OF A MULTI-RANGE THERMO-ELECTRIC WATTMETER

A multi-range wattmeter was constructed using compensated thermal convertors of the type described for the final measuring device, arranged in a system where the voltage drop in the main current circuit did not exceed 0·1 volt, and where the current consumption of the voltage circuit was either 0·2 or 1 mA. The minimum current range was 0·1 amp and the maximum voltage range 500 volts, The accuracy over the frequency range 50 c/s–30 kc/s was comparable with that required by B.S. 89 for precision-grade wattmeters. The system adopted is shown schematically in Fig. 1. A non-inductive 4-terminal resistor R provides at

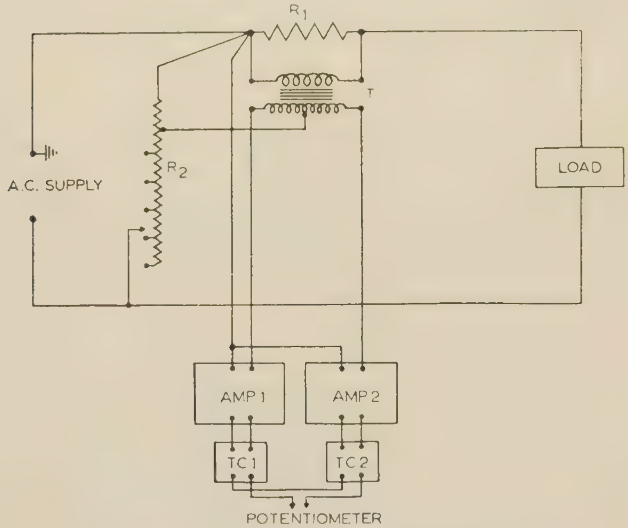


Fig. 1.—Circuit diagram for power measurements.

R_1 Current shunt.
R_2 Tapped voltage-dividing resistor.
T 0·1/1·0 voltage transformer.
AMP 1, AMP 2 Amplifiers.
TC 1, TC 2 Compensated thermal convertor networks.

rated current a voltage of 0·1 volt which is stepped up to 1 volt by the precision voltage transformer T. The centre-tap of this transformer is connected to a point on the voltage divider such that at rated voltage its potential is 0·5 volt above earth. The two voltages fed to the inputs of the two amplifiers are thus the sum and difference respectively of two voltages derived from and proportional to the main-circuit current and voltage. The heater currents of the two thermal convertors are supplied by the outputs of the amplifiers and ideally are proportional to the voltage inputs to the amplifiers. The thermocouples are connected in opposition, and the resultant e.m.f. is thus proportional to the power in the load plus that in the current shunt, together with the transformer losses.

(7) DESCRIPTION AND PERFORMANCE OF THE COMPONENT PARTS OF THE MULTI-RANGE THERMO-ELECTRIC WATTMETER

(7.1) The Compensated Thermal Convertors

One compensated convertor consisted of a convertor having a nichrome heater in series with one having a platinum heater. For the purposes of comparison the other compensated convertor used a nichrome and nickel combination. The compensation and matching of the outputs was carried out experimentally as described earlier, and the detailed performance figures are given in Table 2. The heater current was measured

Table 2

Approximate heater current, I	Approximate output e.m.f., V	Measured values of $V/17{\cdot}920I^2$	
		Ni–Cr + Pt	Ni–Cr + Ni
mA	μV		
7·5	1 000	1·0000	1·0000
10·5	2 000	1·0010	1·0010
13	3 000	1·0015	1·0010
15	4 000	1·0015	1·0010
16·5	5 000	1·0015	1·0010
19	6 500	1·0005	1·0000
20	7 000	1·0000	1·0000
21	8 000	0·9995	0·9995
Total heater-circuit resistance		13 ohms	13 ohms
Change in resistance at 20 mA		+2·5%	+3·5%
Temperature coefficient of output e.m.f. (20–30° C)		+0·02% per deg C	<0·01% per deg C

and maintained constant to 1 part in 10^4, and the output e.m.f. was measured with an uncertainty of 0·5 μV or 2 parts in 10^4, whichever was the greater.

The response time for full output to become established after the application of current was approximately 5 sec. This compares not unfavourably with the limits allowed by B.S. 89: 1954 for the damping of indicating dynamometer wattmeters.

The output of each compensated convertor was measured at a heater current of 20 mA at various frequencies from 50 c/s to 20 kc/s; no change exceeding 2 parts in 10^4 was observed over this range.

(7.2) The Amplifiers

The design of the amplifiers has been described elsewhere.[5] In the present application the normal full output current required is only 20 mA and a small increase in the overall voltage amplification has enabled the input voltage to be reduced from 2 volts to 1 volt. This and other minor modifications in the output stage have resulted in an increased stability margin at frequencies below the working range. The calculated maximum phase shift is now 153° at about 1½ c/s instead of 162° at 2 c/s. The stability conditions at frequencies above the working range are now more favourable, since the highly inductive dynamometer load has been replaced by a resistive thermal convertor. An upward extension of the previous working frequency range was required, and particular attention was therefore paid to the component layout and to the wiring in order to reduce unwanted stray capacitances. Examination of the waveform of the output current obtained with a square-wave input at a frequency of 50 kc/s indicated that the first peak of response with feedback connected occurred at a frequency in excess of 1 Mc/s.

The phase angle between the input voltage and the output current in the working range does not directly affect the accuracy of measurement, the magnitude of the output current being the criterion. Consequently the requirements are (*a*) that the output current of each amplifier shall be proportional to the input voltage and independent of frequency, and (*b*) that the two amplifiers shall give equal outputs for equal inputs. About 40 dB of feedback is provided, and it was found possible to satisfy these requirements to better than ±0·1% for frequencies from 40 c/s to 30 kc/s. The circuit diagram is shown in Fig. 2.

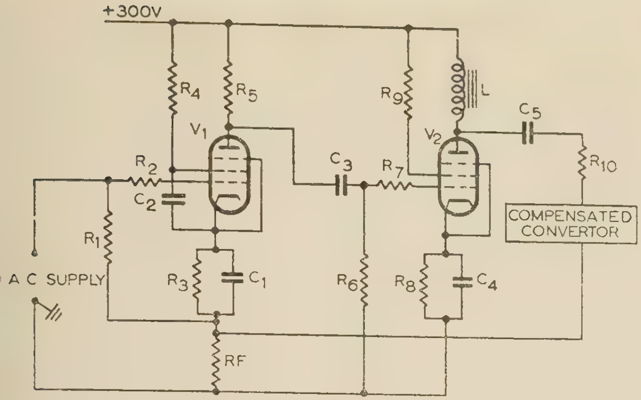


Fig. 2.—Amplifier circuit.

1 MΩ, ½ W		C_1	100 μF, 6 V (electrolytic)
680 Ω, ⅓ W		C_2	0·5 μF, 350 V (paper)
380 Ω, ½ W		C_3	2 μF, 250 V (paper)
100 kΩ, ½ W		C_4	500 μF, 12 V (electrolytic)
39 kΩ, ½ W		C_5	32 μF, 400 V (paper).
150 kΩ, ½ W			
680 Ω, ⅓ W		L	50 H, 80 mA
115 Ω, 1 W			
100 Ω, ½ W		V_1	CV 138
330 Ω, ½ W		V_2	CV 450
50 Ω, non-inductive wire-wound.			

All resistors except RF are high-stability carbon.

A final overall test of the linearity with frequency of an mplifier plus a compensated thermal convertor was made in ach case by measuring the thermocouple output e.m.f. at contant amplifier input voltage. The two output e.m.f.'s were itially made equal for an input voltage of 1·0 volt at a frequency f 1 kc/s by adjustment of one feedback resistor. The frequency as then varied from 40 c/s to 30 kc/s, and no change in output m.f. greater than 3 parts in 10^4 was observed.

(7.3) Voltage Divider and Current Shunts

The method of construction and layout of the voltage divider similar to that previously described.[5] Seven main sections of sistance 2·5, 7·5, 22·5, 47·5, 97·5, 247·5 and 497·5 kilohms e provided which, with a 2·5-kilohm tapping section, gives oltage ranges of 0·5, 1, 5, 10, 20, 50 and 100 volts. The rated rrent consumption for these ranges is thus 0·2 mA. The 5 kilohm tapping section is, however, further subdivided at 00 ohms, and using this as the tapping, ranges up to 500 volts ay be obtained with a current consumption of 1 mA. The re used for the three sections of the highest resistance was e nickel-chromium-aluminium-iron alloy known as Karma, diameter 0·000 6 in and having a resistance of about 2 200 ohms r foot. The use of wire of such high resistance has enabled the erall size and the number of turns for these sections to be duced to about one-third of that previously required. The maining sections used 0·001 or 0·002 in diameter wire of the ckel-chromium-aluminium-copper alloy known as Evanohm. ne temperature coefficient of resistance of all wires used was few parts in 10^6 per deg C.

The differences in time-constant between the tapping and main ctions of the completed divider were measured and found to be ss than 0·05 μH/ohm. Capacitance compensation was applied necessary to reduce the time-constant difference to less than 00 μH/ohm.

The 4-terminal current shunts for providing a voltage proportnal to the circuit current are external to the wattmeter case. ne principles of their design have already been described.[8] ne connection between the potential points of the resistor and e primary leads of the transformer is made through a lowsistance low-inductance connector.

(7.4) The Voltage Transformer

The design of a transformer to operate over a wide band of frequencies involves a balance among conflicting factors. At the low-frequency end, high primary inductance and low winding resistance are necessary, whereas low leakage inductance and capacitance are necessary to maintain a level response at the high-frequency end. The low-frequency performance can be improved by the use of high-permeability magnetic materials, a core of large cross-sectional area and a large number of turns. Both the last two devices, however, increase the leakage inductances and self-capacitances. The leakage inductances can be reduced by interleaving the primary and secondary windings, but only at the expense of increasing the inter-winding capacitances.

The performance demanded of the transformer in the present application was that at any frequency of use the errors should not exceed 0·1% in ratio and 3 min in phase angle. The design and construction of a transformer to comply with these requirements was based on the following considerations:

(*a*) That low-loss magnetic material with an initial relative permeability of the order of 30 000 was available.

(*b*) That the leakage inductances and self-capacitances would be kept to a minimum by the use of single-layer uniformly distributed windings on a toroidal core.

(*c*) That the additional errors at the lowest frequency of use due to the shunting effect of the primary on the 4-terminal current shunt should not exceed 0·1% in ratio and 10 min in angle.

Since the minimum current range of the wattmeter was to be 0·1 amp and the minimum frequency 50 c/s, this final consideration necessitates a minimum value of about 1 henry for the primary inductance of the transformer. It was therefore decided to construct two transformers, one having a primary inductance of about 1 henry at 50 c/s for use from 50 c/s to 5 kc/s, and one having an inductance about 0·1 henry at 500 c/s for use from 500 c/s to 30 kc/s.

Two toroidal cores, 4 in o.d. and 3 in i.d., of high-permeability nickel-iron alloy strip, 0·75 in wide × 0·005 in thick, were tested for initial permeability and power loss at various frequencies. The minimum values of initial permeability obtained were 30 000 at 50 c/s, falling to 10 000 at 10 kc/s. The cores were uniformly wound with primary windings calculated to give the required inductances, a gap of approximately ¾ in being left unwound between the ends of a winding. In order to keep the winding resistance as low as possible, the primary winding was wound first, next to the core. A further advantage is to be obtained from placing the low-voltage winding next to the core: the latter provides an easy path for capacitance currents between turns which are not adjacent, and their effects would be more serious on the high-voltage winding. The inductance and equivalent parallel resistance of the two inductors were then measured in a modified Owen bridge at various frequencies. The values so obtained were in good agreement with the values calculated from the tests on the cores.

The high-frequency errors of a transformer are dependent upon the interwinding capacitances as well as the self-capacitances of the windings. The insulation thickness between the primary and secondary windings therefore has an important effect on the frequency range. Arnold[9] has pointed out that as this thickness is increased from zero the product of the leakage inductance and capacitance decreases to a minimum and then increases, and he gives formulae to evaluate the relationship. For the transformers under consideration it was calculated that the optimum thickness was between 0·04 and 0·09 in. Polyethylene tape and sheet to a total thickness of 0·07 in was used as the insulating medium. For the transformer intended for use at the highest frequencies the experiment of using perforated

polyethylene sheet was tried. It was subsequently found that this reduced the open-circuit inter-winding capacitance by about 20%. No inter-winding screens were used: a double screen is advantageous in concentrating the whole capacitance between two terminals, but it is impossible to use one without increasing the self-capacitances of both windings.

The secondary winding for each transformer was then wound uniformly in a single layer, care being taken to cover the same part of the core occupied by the primary, and to dispose the secondary portions symmetrically with respect to the centre tap. Final binding with 0·002 in polyethylene tape completed the construction. The total leakage inductance and the resonant frequencies with various conditions of connection were then measured. These values and winding data are given in Table 3.

Table 3

	Transformer No. 1	Transformer No. 2
Rating	0·1/1 volt	0·1/1 volt
Primary:		
Wire	4×0·022 in enamelled, in square formation	4×0·022 in enamelled, side by side
No. of turns ..	170	62
D.C. resistance	0·23 ohm	0·08 ohm
Inductance ..	1·0 henry at 50 c/s	0·12 henry at 500 c/s
Secondary:		
Wire	0·0032 in enamelled	0·0108 in enamelled
No. of turns ..	1700	620
Total leakage inductance referred to primary ..	40 μH	5 μH
Resonant frequencies	150, 270 and 500 kc/s	450, 850 and 1500 kc/s

The ratio and phase angle of each transformer were then measured at various frequencies in a bridge circuit by comparison with resistors having known a.c. characteristics up to 20 kc/s. One primary terminal of the transformer was connected to earth, and no changes in value were observed when the primary and secondary connections to the circuit were reversed. The uncertainty in the values of time-constant of the resistors was equivalent to an uncertainty of 2 min in the phase-angle measurements at 20 kc/s. The results obtained are given in Table 4 and refer to zero secondary burden.

Table 4

Primary voltage	Test frequency	Transformer No. 1 True ratio / Nominal ratio	Transformer No. 1 Phase angle	Transformer No. 2 True ratio / Nominal ratio	Transformer No. 2 Phase angle
volt	kc/s		min		min
0·10	0·05	1·0000	+ 2	1·0002	+7
0·05	0·05	1·0000	+ 2	1·0002	+7
0·025	0·05	1·0000	+ 2	1·0002	+8
0·10	0·1	1·0000	+ 1	1·0002	+4
0·10	0·2	1·0000	+ 1	1·0001	+2
0·10	0·5	1·0000	0	1·0001	+1
0·10	1·0	1·0000	0	1·0001	+1
0·10	2·0	1·0000	0	1·0001	0
0·10	5·0	1·0001	+ 1	1·0000	0
0·10	10·0	1·0005	+ 4	1·0001	+1
0·10	15·0	1·0010	+ 5	1·0002	+2
0·10	20·0	1·0015	+10	1·0002	+3

The overall errors which obtain when the transformers are used to measure current, i.e. including the additional errors due to the primary shunting effects, were also measured and found to agree with calculation, being a maximum of 12 min of angle for a current range of 0·1 amp at 50 c/s with transformer No. 1, and at 500 c/s with transformer No. 2.

(8) PERFORMANCE OF THE COMPLETE WATTMETER

The two amplifiers, compensated thermal convertors, voltage divider and voltage transformer were assembled in a well-ventilated metal box measuring approximately 29 × 15 × 10 in deep. Five separate compartments provided electrostatic screening between the units.

Initial overall tests of the wattmeter were made on the 1 amp 50 volt range using transformer No. 2. The errors at full and half load, at unity power factor, did not exceed 0·1%, and the zero-power-factor error did not exceed 0·1% of full load for these conditions in the frequency range 300 c/s to 10 kc/s. Above 10 kc/s, however, the errors increased approximately propor-

Table 5

Current, amp	Circuit power factor	Error % 50 c/s	500 c/s	1 kc/s	5 kc/s	10 kc/s	20 kc/s	30 kc/s
	Range 0·1 amp using transformer No. 1							
0·10	Unity	0·0	0·0	0·0	+0·2			
0·07	Unity	0·0	0·0	0·0	+0·2			
0·05	Unity	0·0	0·0	0·0	+0·2			
0·03	Unity	0·0	0·0	0·0	+0·2			
0·10	Zero lag	+0·3	+0·1	+0·1	+0·2			
0·10	Zero lead	−0·3	−0·1	−0·1	−0·2			
	Range 0·1 amp using transformer No. 2							
0·10	Unity		0·0	0·0	0·0	0·0	0·0	+0·2
0·07	Unity		0·0	0·0	0·0	0·0	0·0	+0·2
0·05	Unity		0·0	0·0	0·0	0·0	0·0	+0·2
0·03	Unity		0·0	0·0	0·0	0·0	0·0	+0·2
0·10	Zero lag		+0·3	0·0	−0·1	−0·1	−0·1	−0·1
0·10	Zero lead		−0·3	−0·1	−0·2	0·0	0·0	−0·1

onally to the square of the frequency, reaching about 1% at)kc/s. In view of the very small individual errors of the various nits comprising the wattmeter, this result was somewhat nexpected. It was deduced that the error was due to the ansformer. When forming part of the wattmeter the necessary onnections alter the distribution of capacitances that obtained hen the transformer was tested as a separate unit. It was not und possible to test the transformer alone under the contions that obtain when it forms part of the working wattmeter. ne inputs leads of the amplifiers also contribute additional pacitance between one primary terminal and each secondary rminal of the transformer. The error due to the latter effect as investigated by connecting capacitance between the primary id secondary, and it was found that about 60 pF caused a ange in error of 1% at 30 kc/s. The input capacitances of the nplifiers were reduced by using the minimum possible length of mi-air-spaced concentric cable, but even so the errors were ll in excess of those required.

The desired high-frequency performance was therefore obtained : a method of compensation. Capacitance was connected in rallel with the heaters of the compensated thermal convertors by-pass the required amount of current. The effect of this -passing was found to be negligible at frequencies below kc/s. The results then obtained on the complete apparatus e given in Table 5, the values referring to rated voltage on any nge.

The zero-power-factor errors at the lowest frequency of use : either transformer do not exceed 0·1% of rated load for rrent ranges greater than 0·5 amp.

The change in error due to a ±10% variation in the test ltage at a constant power was less than 0·1%.

(9) CONCLUSIONS

It has been shown that it is possible by simple means to comnsate commercial thermal convertors so that the net output n.f. is very closely proportional to the square of the heater rrent. The compensation leads to a reduction in the variation output e.m.f. with changes in the ambient temperature. These provements enable the excellent frequency characteristics of ermal convertors to be used to good effect in the precise asurement of a.c. power over a very wide range of frequency. has been found that the combined frequency response of a ermal convertor and an amplifier is constant to better than parts in 10^4 from 40 c/s to 30 kc/s. Whilst no provision has en made for the separate measurement of current and voltage, se results indicate the possibility of constructing an ammeter voltmeter having errors not exceeding 0·1% over this frequency ge.

n the present apparatus the upper frequency limit for the urate measurement of power is fixed by the performance of voltage transformer. Above 10 kc/s it has been necessary to ovide a form of compensation in order to reduce the errors 30 kc/s to about ¼%. Whilst this compensation would be ective in keeping the errors reasonably small at still higher quencies, an increasing dependence would be placed upon it, l further extension of the frequency range by this method is recommended.

(10) EXTENSION OF THE FREQUENCY RANGE

The frequency range has been extended to 100 kc/s and the acitance compensation dispensed with by using a third transner. This was designed for the frequency range 5–100 kc/s. rimary winding of 28 turns of two 0·094 × 0·005 in insulated per tapes connected in parallel was wound on a toroidal core, o.d. and 1½ in i.d., of high-permeability nickel-iron-alloy)·5 in wide × 0·001 in thick. The interwinding insulation was perforated polyethylene 0·07 in thick, and the secondary winding consisted of 280 turns of 0·01 in diameter insulated copper wire. The primary inductance was 10 mH at 5 kc/s, the leakage inductance about 1·5 μH and the lowest resonant frequency about 2 Mc/s. When tested alone over the range 5–100 kc/s the transformer errors did not exceed 0·1%. The wattmeter was tested at unity power factor using this transformer and with the capacitance compensation removed. The overall instrument errors were found not to exceed 0·1% for the range 5–30 kc/s, less than 0·5% at 60 kc/s and less than 1% at 100 kc/s.

It would prove difficult to reduce the time-constants of the voltage divider below their present values, and accurate use of the wattmeter at frequencies above 30 kc/s is restricted thereby to circuits of high power factor.

(11) ACKNOWLEDGEMENTS

The work described has been carried out as part of the research programme of the National Physical Laboratory, and the paper is published by permission of the Director of the Laboratory. The author desires to acknowledge the assistance rendered in the experimental observations by Mr. A. P. Miller.

(12) REFERENCES

(1) MACFADYEN, K. A., and HILL, N. D.: 'A Precision Thermionic A.C. Test Set', *G.E.C. Journal*, 1941, **11**, p. 182.

(2) ARNOLD, A. H. M.: 'Performance Limits in Electrical Instruments', *Proceedings I.E.E.* Paper No. 1128 M, March, 1951 (**98**, Part II, p. 701).

(3) EDMUNDSON, D.: 'Electrical and Magnetic Measurements in an Electrical Engineering Factory', *ibid.*, Paper No. 1703 M, September, 1954 (**102 B**, p. 427).

(4) BOAST, W. B.: 'An Electronic Wattmeter—*Proceedings of the National Electronics Conference* (Chicago), 1952, **8**, p. 716.

(5) ARNOLD, A. H. M., and HILL, J. J.: 'An Audio-Frequency Dynamometer Wattmeter', *Proceedings I.E.E.*, Monograph No. 171 M, March, 1956 (**103 C**, p. 325).

(6) HERMACH, F. L.: 'Thermal Convertors as AC/DC Transfer Standards for Current and Voltage Measurements at Audio Frequencies', *Journal of Research of the National Bureau of Standards*, 1952, **48**, p. 121.

(7) GOFFIN, G., and MARCHAL, G.: 'Étude théorique de l'emploi du thermocouple pour la mesure de précision de la puissance en courant alternatif', *Bulletin de l'Association Suisse des Électriciens*, 1946.

(8) ARNOLD, A. H. M.: 'Alternating-Current Resistance Standards', *Proceedings I.E.E.*, Paper No. 1512 M, June, 1953 (**100**, Pt. II, p. 319).

(9) ARNOLD, A. H. M.: 'The Effect of Capacitance on the Design of Toroidal Current-Transformers', *ibid.*, Paper No. 1020 M, December, 1950 (**97**, Part II, p. 797).

(10) GOODWIN, W. N.: 'The Compensated Thermocouple Ammeter', *Transactions of the American I.E.E.*, 1936, **55**, p. 23.

(13) APPENDIX

The steady-state temperature rise, θ, of an elementary length dx of a uniform conductor heated by an electric current and cooled by conduction to its end terminals and by radiation is governed by the differential equation

$$\frac{d^2\theta}{dx^2} = -\frac{I^2\rho_0(1+\alpha\theta)}{a^2k_0(1+\beta\theta)} + \frac{p\varepsilon\sigma(T^4 - T_0^4)}{ak_0(1+\beta\theta)} \quad . \quad . \quad (6)$$

A rigid solution of this equation would be difficult and the result unnecessarily complicated for the present application to the case of vacuum-enclosed thermal convertors. It is therefore assumed that $\beta\theta$ is small compared to unity and that the total

amount of heat radiated is small compared to that conducted. With these assumptions the following approximations may be made:

(*a*) That $1/(1 + \beta\theta) = (1 - \beta\theta)$ and that this term may be neglected in the second term on the right-hand side of eqn. (6).

(*b*) That $T^4 - T_0^4 = (T_0 + \theta)^4 - T_0^4 \simeq 4\theta T_0^3 + 6\theta^2 T_0^2$. With these approximations eqn. (6) may be rewritten

$$\frac{d^2\theta}{dx^2} = -\frac{I^2\rho_0}{a^2k_0}[1 + (\alpha - \beta)\theta] + \frac{2p\varepsilon\sigma T_0^2}{ak_0}\theta[2T_0 + 3\theta] \quad (6a)$$

Writing $\dfrac{I^2\rho_0}{a^2k_0} = b$ and $\dfrac{p\varepsilon\sigma}{ak_0} = N$, eqn. (6*a*) simplifies to

$$\frac{d^2\theta}{dx^2} = -b[1 + (\alpha - \beta)\theta] + 2NT_0^2\theta[2T_0 + 3\theta] \quad (6b)$$

If the resistivity and thermal conductivity of the conductor are invariable with temperature and there are no radiation losses eqn. (6*b*) reduces to

$$\frac{d^2\theta}{dx^2} = -b \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6c)$$

Considering the origin to be at the mid-point, the boundary conditions are $d\theta/dx = 0$ at $x = 0$ and $\theta = 0$ at $x = \pm l$.

On integration
$$\theta = \frac{b}{2}(l^2 - x^2) \quad . \quad . \quad . \quad . \quad . \quad (7)$$

and the mid-point temperature rise is
$$\theta_M = \frac{bl^2}{2} \quad . \quad . \quad . \quad (8)$$

If each of the disturbing terms which cause eqn. (6*b*) to differ from eqn. (6*c*) are small, sufficient accuracy will be obtained by substituting into (6*b*) the value of θ obtained from eqn. (7) and solving the resulting equation. With this substitution eqn. (6*b*) becomes

$$\frac{d^2\theta}{dx^2} = -b\left[1 + (\alpha - \beta)\frac{b}{2}(l^2 - x^2)\right] + 2NT_0^2\left\{\frac{b}{2}(l^2 - x^2)\left[2T_0 + 3\frac{b}{2}(l^2 - x^2)\right]\right\} \quad (9)$$

Direct integration of (9) and application of the boundary conditions gives the mid-point temperature rise

$$\theta_M = \frac{bl^2}{2}\left[1 + \frac{5}{6}\frac{bl^2}{2}(\alpha - \beta) - \frac{5}{3}NT_0^3l^2 - \frac{11}{5}NT_0^2l^2\frac{bl^2}{2}\right] \quad (10)$$

The terms in eqn. (10) are the same as the relevant terms given in Hermach's[6] equation for the low-frequency error of a thermal convertor, which was developed in a similar manner. When eqn. (10) is evaluated for vacuum-enclosed thermal convertors having heaters made from the usual resistance alloys, it is found that the disturbance caused by the term in β is usually the greatest single cause of error and is frequently greater than the sum of the remaining terms. In some cases it may be large enough to prevent sufficient accuracy being obtained. A closer approximation would therefore be obtained by including the β term from the beginning and then using the value of θ so found for substitution into (6*b*).

Consider, then, the equation

$$\frac{d^2\theta}{dx^2} = -b(1 - \beta\theta) \quad . \quad . \quad . \quad . \quad (11)$$

Integration and application of the boundary conditions gives

$$\theta = \frac{1}{\beta}\left[1 - \frac{\cosh\sqrt{(b\beta)}x}{\cosh\sqrt{(b\beta)}l}\right] \quad . \quad . \quad . \quad . \quad (12)$$

and the mid-point temperature-rise

$$\theta_M = \frac{bl^2}{2} - \frac{5}{24}\beta b^2l^4 + \frac{61}{720}\beta^2b^3l^6 - \ldots \quad . \quad (13)$$

A slight adjustment in the numerical coefficient of the term in l^6 and neglect of terms of higher powers gives

$$\theta_M \simeq \frac{bl^2}{2}\left[1 - \frac{5}{6}\beta\frac{bl^2}{2}\left(1 - \frac{5}{6}\beta\frac{bl^2}{2}\right)\right] \quad . \quad . \quad (14)$$

It is to be noted that the term in β in eqn. (10) has been modified by being multiplied by the factor $\left(1 - \frac{5}{6}\beta\frac{bl^2}{2}\right)$. Thus by taking the first few terms in the expansion of eqn. (12) and with the same slight numerical adjustment we may write

$$\theta = \frac{b}{2}(l^2 - x^2)\left(1 - \frac{5}{6}\beta\frac{bl^2}{2}\right)$$

and use this instead of $\theta = b/2(l^2 - x^2)$ for substitution into eqn. (6*b*), which then becomes

$$\frac{d^2\theta}{dx^2} = -b\left[1 + (\alpha - \beta)\frac{b}{2}(l^2 - x^2)\left(1 - \frac{5}{6}\beta\frac{bl^2}{2}\right)\right] + 2NT_0^2\frac{b}{2}(l^2 - x^2)\left(1 - \frac{5}{6}\beta\frac{bl^2}{2}\right) \times \left[2T_0 + 3\frac{b}{2}(l^2 - x^2)\left(1 - \frac{5}{6}\beta\frac{bl^2}{2}\right)\right] \quad (15)$$

Integrating and applying the boundary conditions then gives the mid-point temperature rise

$$\theta_M = \frac{bl^2}{2}\left\{1 - \frac{5}{3}NT_0^3l^2\left(1 - \frac{5}{6}\beta\frac{bl^2}{2}\right) + \left[\frac{5}{6}\alpha - \frac{5}{6}\beta - \frac{11}{5}NT_0^2l^2\left(1 - \frac{5}{6}\beta\frac{bl^2}{2}\right)\right]\left(1 - \frac{5}{6}\beta\frac{bl^2}{2}\right)\frac{bl^2}{2}\right\} \quad . \quad . \quad . \quad . \quad (16)$$

Eqn. (16) has allowed for the disturbing effects due to temperature coefficients of resistivity and thermal conductivity and to radiation. In a complete thermal convertor a further disturbance is caused by the attached thermocouple, and Goodwin[10] has developed equations to evaluate this effect. These show that the mid-point temperature rise has to be multiplied by a factor which is approximately equal to

$$\frac{a_hk_hl_c}{a_hk_hl_c + a_ck_cl_h}$$

where the suffixes h and c refer to the heater and thermocouple wires respectively. Assuming that we may put $k_h\rho_h \simeq k_c\rho_c$ then the above factor reduces to $R_c/(R_c + R_h)$, where R_h and R_c are the resistances of the heater and thermocouple wires respectively. Eqn. (16) must therefore be multiplied by this factor, and the final expression, reintroducing the original symbols, can be written as

$$\theta_M = \frac{R_c}{R_c + R_h} \times \frac{I^2\rho_0l^2}{2a^2k_0}\left[1 + \frac{5}{6}\alpha\frac{I^2\rho_0l^2}{2a^2k_0}\left(1 - \frac{5}{6}\beta\frac{I^2\rho_0l^2}{2a^2k_0}\right) - \frac{5}{6}\beta\frac{I^2\rho_0l^2}{2a^2k_0}\left(1 - \frac{5}{6}\beta\frac{I^2\rho_0l^2}{2a^2k_0}\right) - \frac{5}{3}\frac{p\varepsilon\sigma T_0^3l^2}{ak_0}\left(1 - \frac{5}{6}\beta\frac{I^2\rho_0l^2}{2a^2k_0}\right) - \frac{11}{5}\frac{p\varepsilon\sigma T_0^2l^2}{ak_0}\frac{I^2\rho_0l^2}{2a^2k_0}\left(1 - \frac{5}{6}\beta\frac{I^2\rho_0l^2}{2a^2k_0}\right)^2\right]$$

1.372.823

The Institution of Electrical Engineers
Paper No. 2468 R
Jan. 1958

A NOTE ON THE FOURIER SERIES REPRESENTATION OF THE DISPERSION CURVES FOR CIRCULAR IRIS-LOADED WAVEGUIDES

By P. N. ROBSON, B.A., Associate Member.

(*The paper was first received* 2nd *August, and in revised form* 15*th October*, 1957.)

SUMMARY

The dispersion curve for the lowest pass band in a uniform corru-ed circular waveguide, such as is used in travelling-wave linear celerators, may be expressed as a Fourier series of period 2π/(corru-ion pitch). Under the slowest conditions of convergence in the veguides investigated, an accuracy of ± 1 part in 15 000 in the ermination of frequency corresponding to a given phase velocity ı be obtained by using only six terms of this series. For the majority structures considered, three terms only are required. An empirical ression for the dispersion curve can thus be obtained from only a all number of experimental results, and is of considerable use in design of suitable slow-wave structures for accelerators. The ation of the dispersion curve can be readily differentiated to give group velocity at all points within the pass band.

The validity of the method is illustrated using a selection of experi-ntal results which have been obtained in the course of an extensive gramme of measurements on corrugated slow-wave structures for ear accelerators operating at wavelengths around 10 cm.

LIST OF PRINCIPAL SYMBOLS

a = Radius of iris hole in circular waveguide.
b = Radius of outer wall in circular waveguide.
c = Velocity of light in vacuo.
D = Corrugation spacing.
$(D - d)$ = Thickness of corrugation iris.
f = Frequency.
v_g = Group velocity.
v_p = Phase velocity.
β = Propagation coefficient = $2\pi/\lambda_g$.
λ_0 = Free-space wavelength.
λ_g = Guide wavelength.
ω = Angular frequency.

(1) INTRODUCTION

The slow-wave structure used in the majority of travelling-ve electron accelerators is the iris-loaded circular waveguide, pagating the circularly symmetric TM mode, shown in tion in Fig. 1. In such accelerators the electrons and wave

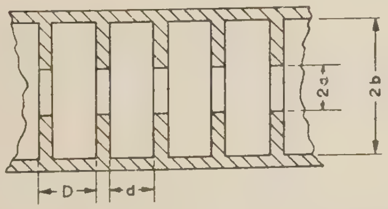


Fig. 1.—Section of iris-loaded circular waveguide.

ve to travel in synchronism over many wavelengths for ctive acceleration, and it follows that the phase velocity of wave must be very closely matched to the electron velocity.

If the electron velocity at any point along the accelerator is known, it remains to decide on the dimensions of the corrugated slow-wave structure to give a phase velocity equal to this electron velocity. The corrugation pitch D is determined from considerations of cut-off wavelength and attenuation,[1,2] the iris-hole radius a is fixed by the desired series impedance, and the iris thickness $(D - d)$ is sufficient to give good mechanical rigidity and to prevent excessively high field strengths along the perimeter of the iris hole. Typical values for these dimensions are:

$$D/\lambda_0 = 0{\cdot}2 \qquad a/\lambda_0 = 0{\cdot}13 \qquad d/\lambda_0 = 0{\cdot}175$$

It is convenient to normalize these dimensions with respect to the free-space wavelength λ_0. When suitable values for a, d, and D have been chosen, it remains to determine b/λ_0 to give the desired phase velocity v_p. The problem of relating the above dimensions to phase velocity has formed the subject of several papers.[2–5] The only analytical methods that give results sufficiently accurate for design purposes require considerable numerical work. In practice, a suitable structure is made and the variation of propagation coefficient with frequency is measured experimentally. From such information, obtained for several structures of slightly differing dimensions, it is possible by suitable scaling and interpolation to obtain the dimensions of a structure to provide any likely propagation coefficient at the design frequency.

The group velocity v_g of this type of corrugated structure is very low; a typical value being $c/30$.

Since $v_g = d\omega/d\beta$ and $v_p = \omega/\beta$ it follows that

$$\frac{dv_p}{v_p} = \frac{d\omega}{\omega}\left(1 - \frac{\omega}{\beta}\frac{d\beta}{d\omega}\right)$$
$$= \frac{d\omega}{\omega}\left(1 - \frac{v_p}{v_g}\right)$$

Now $v_p/v_g \simeq 30$ and since $|d\omega/\omega| \simeq |da/a| \simeq |db/b|$

$$\left|\frac{dv_p}{v_p}\right| \simeq -30\left|\frac{da}{a}\right| \simeq -30\left|\frac{db}{b}\right|$$

Thus a fractional change in either a or b produces approximately a 30-fold fractional change in phase velocity. It is necessary, therefore, to be able to determine the dispersion characteristics, i.e. frequency versus propagation coefficient, to a high degree of accuracy, say to 1 part in 10 000 for a given propagation coefficient β. For the structures considered later, an extremely simple semi-empirical expression of high accuracy can be obtained as outlined in the next Section.

(2) THEORY

It is known that, for wave propagation in any lossless periodic waveguide structure, the frequency in a particular pass band is an even periodic function of the propagation coefficient.[6] For a corrugated waveguide propagating the circularly symmetric TM mode, the form of a typical frequency/propagation-coefficient

ritten contributions on papers published without being read at meetings are ed for consideration with a view to publication.
r. Robson, who was formerly with the Metropolitan-Vickers Electrical Co. Ltd., ow at the University of Sheffield.

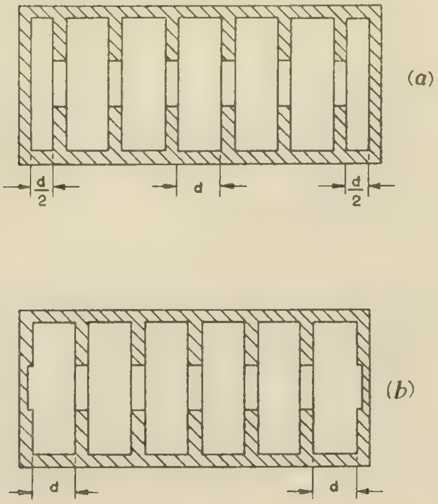


Fig. 2

(*a*) Cavity made up from length of corrugated circular waveguide and terminated by two half-sections.

(*b*) Cavity made up from length of corrugated circular waveguide and terminated by two full-sections.

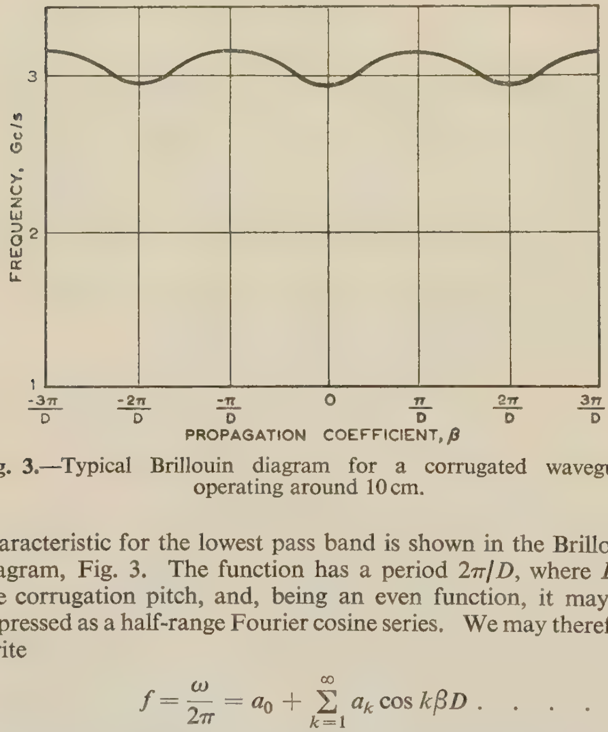


Fig. 3.—Typical Brillouin diagram for a corrugated waveguide operating around 10 cm.

characteristic for the lowest pass band is shown in the Brillouin diagram, Fig. 3. The function has a period $2\pi/D$, where D is the corrugation pitch, and, being an even function, it may be expressed as a half-range Fourier cosine series. We may therefore write

$$f = \frac{\omega}{2\pi} = a_0 + \sum_{k=1}^{\infty} a_k \cos k\beta D \quad . \quad . \quad . \quad . \quad (1)$$

where β is the propagation coefficient, f is the frequency, and the coefficients $a_0 \ldots a_k$ are determined by the geometry of the guide.

If the coefficients in eqn. (1) converge rapidly, only a few terms suffice to give an accurate expression for the frequency in terms of the propagation coefficient. The coefficients a_0, a_1, a_2, etc., can be obtained for any given guide from a few measurements of frequency and guide wavelength made at suitable points in the pass band. It is shown later that the convergence of series (1) is extremely rapid in all the waveguides that were measured; in no instances are more than six terms required to give an accuracy better than 1 part in 15 000, and in general only three or four terms are required.

The measurement of certain parameters associated with corrugated structures, namely series impedance, shunt impedance and attenuation, require a knowledge of the variation of group velocity within the pass band.[7] This is readily obtainable by differentiating eqn. (1):

$$v_g = \frac{d\omega}{d\beta} = -2\pi D \sum_{k=1}^{\infty} k a_k \sin k\beta D \quad . \quad . \quad . \quad . \quad (2)$$

Although the convergence of eqn. (2) is not so rapid as that of (1), it may provide a more accurate method of determining v_g than by numerical or graphical differentiation of an experimental dispersion curve.[7]

(3) TECHNIQUE

The determination of the coefficients $a_0 \ldots a_k$ is simplified considerably if values of frequency, $f_0 \ldots f_t$ ($t \geqslant k$), can be obtained for a number of equally spaced values of β within one half-period.

A cavity is made up[7] of a length of corrugated waveguide terminated at either end by reflecting planes, such planes being placed either at the mid-section of a corrugation or of an iris in order to preserve the symmetry. Fig. 2 shows the former and latter methods of termination for a cavity having six corrugations.

If the cavity is n corrugations long, the possible modes of oscillation are given by

$$m\frac{\lambda_g}{2} = nD$$

where m has the values 0, 1, 2 ... n for a termination as shown in Fig. 2(*a*), or 0, 1, 2 ... ($n - 1$) when the cavity is terminated as in Fig. 2(*b*).

Therefore
$$\lambda_g = \frac{2nD}{m} = \frac{2\pi}{\beta}$$

whence
$$\beta = \frac{\pi m}{nD} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

When $m = 0$, there is no phase change between consecutive corrugations and this field configuration is designated the 0-mode; when $m = n$, the phase change per corrugation is π radians and the corresponding configuration is now called the π-mode. The absence of the π-mode for the lowest pass band when the cavity is terminated as in Fig. 2(*b*) has been dealt with by several authors.[4,8] In view of this, all investigations of the lowest pass band were done with cavities terminated as in Fig. 2(*a*), and therefore having the π-mode present.

If, for example, $n = 6$, the cavity will resonate at the seven frequencies $f_0, f_1 \ldots f_6$ corresponding to $m = 0, 1 \ldots 5, 6$. It can be shown[9] that under these conditions the coefficients, as far as a_6, may be determined from these values of f by the following relationships:

$$12a_0 = (f_0 + 2f_1 + 2f_2 + 2f_3 + 2f_4 + 2f_5 + f_6) \quad (5a)$$
$$6a_1 = (f_0 + f_2 - f_4 - f_6) + \sqrt{3}(f_1 - f_5) \quad . \quad . \quad (5b)$$
$$6a_2 = (f_0 + f_1 + f_5 + f_6) - (f_2 + f_4 + 2f_3) \quad . \quad (5c)$$
$$6a_3 = (f_0 - f_6 + 2f_4 - 2f_2) \quad . \quad . \quad . \quad . \quad . \quad . \quad (5d)$$
$$6a_4 = (f_0 + f_6 + 2f_3) - (f_1 + f_2 + f_4 + f_5) \quad . \quad (5e)$$
$$6a_5 = (f_0 + f_2 - f_4 - f_6) + \sqrt{3}(f_5 - f_1) \quad . \quad . \quad (5f)$$
$$12a_6 = (f_0 - 2f_1 + 2f_2 - 2f_3 + 2f_4 - 2f_5 + f_6) \quad (5g)$$

The resonant frequencies were determined by measuring the cavity responses around each resonance and plotting the corre-

onding Q-curves.[7] The frequency of the klystron local cillator was measured using a high-Q, H_{011} cavity wavemeter. ne experimental error in measuring the frequency for any ven β was estimated to be about ± 1 part in 15000. There is no int, therefore, in trying to evaluate coefficients smaller than /15000. In the majority of cases considered, a_3 and higher efficients can be neglected on this account. When only three rms are required to give sufficient accuracy,

$$4a_0 = (f_0 + f_6 + 2f_3) \quad . \quad . \quad . \quad . \quad . \quad (5h)$$

$$2a_1 = (f_0 - f_6) \quad . \quad . \quad . \quad . \quad . \quad . \quad (5i)$$

$$4a_2 = (f_0 - 2f_3 + f_6) \quad . \quad . \quad . \quad . \quad . \quad (5j)$$

nere f_0, f_3 and f_6 are the frequencies of the 0, $\pi/2$ and π modes spectively. Eqns. (5a)–(5c), however, present a determination these three coefficients, if sufficient data are available, that is ss sensitive to observational error than when only three freencies are used. The coefficients a_0, a_1 and a_2 under these nditions represent the best fit in accordance with the criterion least squares.

(4) RESULTS

(4.1) Dispersion Curve Equations

The dispersion curve for a corrugated waveguide having = 2·5 cm, $a = 0{\cdot}9948$ cm, $b = 3{\cdot}9310$ cm and $(D - d) =$ 5560 cm was evaluated from the resonant frequencies of the $\pi/2$ and π modes using equations (5h)–(5j). The resulting ries is given below, together with Table 1, comparing the

Table 1

$\lambda_g = \frac{2\pi}{\beta}$	f (calculated)	f (experimental)	Difference
cm	Mc/s	Mc/s	Mc/s
5	2991·65	2991·65	0·00
6·25	2988·80	2988·65	−0·15
7·5	2984·30	2984·30	0·00
8·33	2981·42	2981·35	−0·07
10	2976·75	2976·75	0·00
12·5	2972·02	2972·00	−0·02
15	2969·06	2969·05	−0·01
∞	2961·20	2961·20	0·00

perimentally obtained resonant frequencies for intermediate odes in the pass band with those predicted by eqn. (6). It l be seen that agreement is extremely good:

$$f = 2976{\cdot}59 - 15{\cdot}22 \cos \beta D - 0{\cdot}165 \cos 2\beta D \quad . \quad . \quad (6)$$

A selection of results for corrugated waveguides of varying nensions is given in Table 2.

The convergence of the series for waveguides 1–3 is extremely pid; all higher-order terms that are omitted were less than the perimental error of about $\pm 0{\cdot}2$ Mc/s. With waveguide 4, having a smaller corrugation pitch of only 1 cm, the convergence is much less rapid.

(4.2) Group Velocity Measurements

Eqn. (2) gives

$$v_g = -2\pi D \sum_{k=1}^{\infty} k a_k \sin k\beta D$$

The convergence of this series is slower than that of the series in eqn. (1) since the general coefficient a_k is now multiplied by k.

The magnitude of the group velocity is determined largely by the first term a_1 of series (2), particularly in the centre of the pass band ($\beta D \simeq \pi/2$), which is usually the section of interest. The error due to omitting coefficients higher than a_{k-1} is approximately ka_k.

Therefore

$$\text{Percentage error in } v_g = \frac{ka_k}{a_1} 100\%$$

If it is desired to limit this error to 1%, the series must be evaluated as far as the term in a_{k-1}, where k is given by

$$ka_k = \frac{a_1}{100}$$

Reference to earlier results shows in general that, to obtain 1% accuracy for waveguides 1–3, two terms only are required, whereas five terms are required for waveguide 4 (see Table 2). In general, if the expansion relating β and f is accurate to ± 1 part in 15000, the group velocity obtained by differentiating this equation is accurate to better than 1% at the centre of the pass band. The accuracy tends to fall off at either end of the pass band.

(5) CONCLUSIONS

It would appear that for the range of corrugated circular waveguides used in linear accelerators, the dispersion characteristics, i.e. frequency/propagation-coefficient, may be expressed with considerable accuracy as the first few terms of a Fourier series. For a waveguide having ten corrugations per wavelength, about six terms are required in the expansion, but for a waveguide with five or less corrugations per wavelength, only three terms are required. The convergence of these series is such that they may be differentiated to obtain a further expression for the group velocity, in error by less than 1%.

The technique described presents an accurate and rapid way of determining the dispersion curve. In the majority of structures used for travelling-wave linear accelerators, only three point frequencies, corresponding to the 0, $\pi/2$ and π modes respectively, are needed to determine the equation of this curve.

The majority of slow-wave structures used in travelling-wave tubes and backward-wave oscillators are much less dispersive than those already described, and consequently the coefficients of eqn. (1) are not likely to decrease so rapidly as in the case of

Table 2

No.	a	b	D	$(D-d)$	Dispersion relationship:
	cm	cm	cm	cm	Mc/s
1	1·50	4·006	2·5	0·25	$3018{\cdot}13 - 104{\cdot}13 \cos \beta D - 2{\cdot}67 \cos 2\beta D$
2*	1·25	3·949	2·5	0·25	$3001{\cdot}52 - 60{\cdot}23 \cos \beta D - 1{\cdot}075 \cos 2\beta D$
3	1·30	3·9725	2	0·25	$3017{\cdot}79 - 80{\cdot}55 \cos \beta D - 3{\cdot}42 \cos 2\beta D$
4	1·3	4·0129	1	0·25	$3091{\cdot}89 - 108{\cdot}96 \cos \beta D - 18{\cdot}91 \cos 2\beta D - 3{\cdot}58 \cos 3\beta D - 1{\cdot}11 \cos 4\beta D - 0{\cdot}19 \cos 5\beta D$

* Results taken from data published by Grosjean.[10]

accelerator structures. However, this method may still provide a convenient way of presenting the dispersion characteristics for such structures.

(6) ACKNOWLEDGMENTS

The author wishes to thank Dr. Willis Jackson, Director of Research and Education, and Mr. B. G. Churcher, Manager of the Research Department, Metropolitan-Vickers Electrical Co. Ltd., for permission to publish the paper.

(7) REFERENCES

(1) CHODOROW, M., *et al.*: 'Stanford High Energy Linear Electron Accelerator (Mark III)', *Review of Scientific Instruments*, 1955, **26**, p. 134.

(2) WALKINSHAW, W.: 'Theoretical Design of Linear Accelerator for Electrons', *Proceedings of the Physical Society*, 1948, **61**, p. 246.

(3) CHU, E. L.: 'The Theory of Linear Accelerators', M.L. Report No. 140, Microwave Laboratory, Stanford University, 1951.

(4) GROSJEAN, C. C.: 'On the Theory of Circularly Symmetric TM Waves in Infinite Iris-loaded Guides', *Nuovo Cimento*, 1955, **1**, p. 427.
'Theory of Circularly Symmetric Standing TM Waves in Terminated Iris-loaded Guides', *ibid.*, 1955, **2**, p. 11.

(5) WALKINSHAW, W. W., and BELL, J. S.: 'Review of Theory of Metal Loaded Linear Accelerator', A.E.R.E. Report G/R 675, 1951.

(6) SLATER, J. C.: 'Microwave Electronics' (Van Nostrand, 1950), Chap. 8.

(7) AIKIN, A. W.: 'Measurements in Travelling-Wave Structures', *Wireless Engineer*, 1955, **32**, p. 230.

(8) LAWTON, E. J.: 'The Choice of Operating Mode for Standing Wave Type Linear Accelerators for Electrons', *Journal of Applied Physics*, 1948, **19**, p. 534.

(9) WHITTAKER, E. T., and ROBINSON, G.: 'The Calculus of Observations' (Blackie and Sons, 1919), Chap. 10.

(10) GROSJEAN, C. C., and VANHUYSE, V. J.: 'Experimental Verification of a Frequency Equation for Corrugated Waveguides', *Nuovo Cimento*, 1955, **1**, p. 193.

1.37 : 621.372.8 : 621.315.61

The Institution of Electrical Engineers
Paper No. 2480 R
Jan. 1958
©

HE USE OF DIELECTRIC MATERIALS TO ENHANCE THE REFLECTIVITY OF A SURFACE AT MICROWAVE FREQUENCIES

By G. B. WALKER, M.A., Ph.D., and J. T. HYMAN, B.Sc.(Eng.), Graduate.

(*The paper was first received* 29*th June, and in revised form* 24*th October*, 1957.)

SUMMARY

The padding of a metal surface by one or more layers of a dielectric .terial is examined, and it is shown that an improvement in reflectivity ı be obtained provided that the loss factor (tan δ) of the dielectric mall enough. For most materials the gain is approximately propor-nal to $\sqrt{\epsilon}/\tan\delta$, where ϵ is the dielectric constant.

Since the reflectivity of a metal diminishes as the frequency is reased, it is to be expected that dielectric padding will become more vantageous at the higher frequencies which can now be generated. r known materials, no improvement is to be expected at frequencies .ch below 3 000 Mc/s, but at that frequency a gain of two was served for a ceramic material consisting mainly of titanium dioxide.

By laminating a dielectric it is possible to reduce the effective loss tor. A reduction in ϵ must occur, but an overall increase in $/\tan\delta$ can be obtained. In the case of titania, laminating can give urther improvement by a factor of three, and hence at 3 000 Mc/s it heoretically possible to produce a surface which is six times as ective as a metal. At higher frequencies even greater gains may possible.

LIST OF PRINCIPAL SYMBOLS

, ϵ_0 = Permeability and permittivity, respectively, of free space.
ϵ = Dielectric constant.
Z_A = Characteristic wave impedance in an air filled region.
Z_D = Characteristic wave impedance in a dielectric filled region.
Z_M = Characteristic wave impedance of a metal.
n δ = Loss factor of dielectric material.
f = Frequency.
f_c = Cut-off frequency in air filled guide.

(1) INTRODUCTION

n all microwave devices some power is absorbed by the walls taining the electromagnetic fields, and in many cases the ciency of the device is very largely determined by this wall loss. considerable amount of work has been done in recent years connection with the utilization of metal walls, but there is reason to suppose that a different order of surface resistance ı be attained. Moreover, in a metal wall the effective surface stance increases with frequency, owing to skin effect, and so l loss becomes a more serious problem at the higher fre-encies which can now be generated.*

There is no known material which in any way approaches a tal in reflective power, but the transformer property of a arter-wave thickness of a dielectric can be used to prepare ficially a surface which is highly reflecting. Much use has been made in optics of dielectric layers, especially for the converse problem of producing a perfectly absorbing surface, but the technique does not appear to have received a comparable amount of consideration at microwave frequencies, and it is not generally known that with dielectric materials available at present it is, in fact, possible to produce a surface which is more reflective than a metal.

In this paper an attempt is first made to set down a criterion which will decide whether any particular dielectric material is suitable for this purpose, and subsequently an experiment is described which demonstrates the advantage to be gained.

(2) SPECIFICATION OF WALL LOSS

To compare the quality of different surfaces, the ratio of absorbed power to incident power may be taken. Factors such as the shape of the surface and the electromagnetic field pattern are important, but for most purposes a satisfactory comparison can be made by evaluating this ratio for a plane surface on to which a wave is incident in the direction of the normal. For waves in free space and for all propagating modes in a waveguide the ratio of the transverse E and H field components is constant over a wavefront and is known as the wave impedance. Since the power flow depends only on the transverse field components it is convenient to express the absorption ratio r in terms of the characteristic impedance Z_A of the incident wave and the impedance Z_1 existing at the absorbing surface. It can be shown that

$$r = \frac{4\mathscr{R}Z_1\left|\frac{Z_A}{Z_1 + Z_A}\right|^2}{\mathscr{R}Z_A} \quad . \quad . \quad . \quad . \quad . \quad (1)$$

The ratio tends to zero if Z_1 is very large or very small compared with Z_A, and hence a perfect reflection occurs if the wall acts either as an open-circuit or as a short-circuit. For reasons to be given later the open-circuit case is not of practical interest, and in what follows only the short-circuit case will be considered. For this the absorption ratio reduces to $4\mathscr{R}Z_1/\mathscr{R}Z_A$. It may be noted that this approximation is equivalent to the one it is customary to use in calculating wall loss, the field outside the wall being calculated on the assumption that no wall loss occurs, the wall loss being subsequently deduced from a knowledge of the magnetic field strength at the wall.

(3) REFLECTION FROM LOSS-FREE QUARTER-WAVE DIELECTRIC LAYERS

As will be shown later, the improvement in reflection to be obtained by the use of quarter-wave laminations of dielectric material (dielectric padding) is almost entirely determined by the loss factor of the material. It is instructive, however, to consider first the ideal case of loss-free material.

The effect of a single quarter-wave thickness of dielectric is to invert the impedance. Thus, if a sheet of characteristic impedance Z_D is placed in contact with a metal surface of impedance Z_M, the impedance at the outer face of the dielectric is Z_D^2/Z_M.

* The propagation of an H_{01} mode in a circular waveguide might seem to be an ption. In this case, however, the diminution of attenuation with frequency is due e change in field pattern which over-compensates the increase in absorptive power e wall. For a given power flow in the guide, the magnetic field strength at the diminishes in proportion to the frequency, whereas the absorptive quality of the increases in proportion to the square root of the frequency.

itten contributions on papers published without being read at meetings are ed for consideration with a view to publication.

r. Walker is at Queen Mary College, University of London.
r. Hyman, who was formerly at that College, is now at the United Kingdom m. Energy Research Establishment.

If two dielectric layers are used, having impedances Z_{D1} and Z_{D2}, the latter referring to the material in contact with the metal, the impedance at the outer face is $(Z_{D1}/Z_{D2})^2 Z_M$. The process can be continued for multi-layers, each additional layer causing a further inversion.

To obtain an enhanced reflection it is clearly desirable to use alternate layers with the greatest possible difference in impedance. For non-magnetic substances the highest wave impedance is that of free space, Z_A, and hence the greatest effect is to be obtained by the use of dielectric layers separated by air spaces.

It may be remarked that, by having an odd number of quarter-wave laminations, a high-impedance or open-circuit boundary can be produced. It is easy to show, however, that unless a material having wave impedance greater than air is placed in contact with the metal, no improvement in reflection can result. It is possible that a ferromagnetic material could be found which would give an enhanced reflection in this way, but this question will not be discussed here.

(4) REFLECTION FROM A DOUBLE LAYER

In Fig. 1 a sheet of dielectric material of thickness d is placed at a distance D from a metal surface. The air space may be

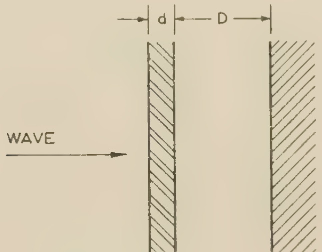


Fig. 1.—Dielectric sheet placed in front of a metal surface.

regarded as loss-free, but in the dielectric the propagation coefficient is complex, having a real part, or attenuation coefficient, α, and an imaginary part, or phase-change coefficient, $j\beta$. Taking D as one-quarter of the wavelength in the air region, the wave impedance at the surface of the dielectric nearest to the metal is Z_A^2/Z_M (as mentioned in the previous Section), and hence, by the well-known impedance transformation, the wave impedance at the outer surface of the dielectric is

$$Z = Z_D \frac{Z_A^2/Z_M + Z_D \tanh \gamma d}{Z_D + (Z_A^2/Z_M) \tanh \gamma d} \quad . \quad . \quad . \quad (2)$$

It can be shown that $\mathscr{R}Z$ is a minimum when d is a quarter of the wavelength in the dielectric region, i.e. for $\beta d = \pi/2$. Using this value of d, it can be shown that $\alpha d \simeq (\pi/4) \tan \delta$, where $\tan \delta$ is the loss factor of the material, which for low-loss materials is less than 10^{-3}. It is therefore permissible to replace $\tanh \gamma d$ by $1/(\pi/4) \tan \delta$. In the denominator of the expression for Z given in (2) the term Z_D is negligible compared with $(Z_A^2/Z_M) \tanh \gamma d$, and hence one may write, with an error of less than 0·1%,

$$Z = (Z_D/Z_A)^2 Z_M + \frac{\pi}{4} Z_D \tan \delta \quad . \quad . \quad . \quad (3)$$

A physical meaning can be given to the two terms which define the surface impedance Z. The first, namely $(Z_D/Z_A)^2 Z_M$, takes account of the metal wall and shows that the loss in the metal wall is reduced by the factor $(Z_D/Z_A)^2$. The second term, namely $(\pi/4) Z_D \tan \delta$, depends entirely on the dielectric and accounts for the energy absorbed there.

(5) REQUIREMENT FOR ENHANCED REFLECTION

The condition that the reflection at the outer surface of the dielectric is better than for a metal surface without padding is that $\mathscr{R}Z < \mathscr{R}Z_M$, and hence by (3) it is necessary that

$$\tan \delta < (4/\pi Z_D)[1 - (Z_D/Z_A)^2] \mathscr{R}Z_M \quad . \quad . \quad (4)$$

It is interesting to note that, if infinitely many layers of dielectric are used, the impedance at the outer surface, as may be deduced from (3), is

$$Z = \frac{\pi}{4} \tan \delta Z_D \left[1 + \left(\frac{Z_D}{Z_A}\right)^2 + \left(\frac{Z_D}{Z_A}\right)^4 + \ldots\right]$$

$$= \frac{\pi}{4} \tan \delta Z_D \{1/[1 - (Z_D/Z_A)^2]\} \quad . \quad . \quad . \quad . \quad . \quad (5)$$

and hence the inequality (4) is also the condition that the reflectivity of an infinite stack of quarter-wave dielectric layers exceeds that of a metal. It is therefore a perfectly general criterion.

To illustrate the importance of the dielectric constant it is helpful to evaluate Z_D and Z_A for an unbounded plane wave, writing $(Z_D/Z_A)^2 = 1/\epsilon$ and $Z_D = \sqrt{(\mu/\epsilon_0 \epsilon)} = 120\pi/\sqrt{\epsilon}$. Eqn. (3) then becomes

$$Z = \frac{1}{\epsilon} Z_M + \frac{30\pi^2 \tan \delta}{\sqrt{\epsilon}} \quad . \quad . \quad . \quad . \quad . \quad (6)$$

and the inequality (4) becomes, after rearrangement,

$$\tan \delta \frac{\sqrt{\epsilon}}{\epsilon - 1} < \frac{\mathscr{R}Z_M}{30\pi^2} \quad . \quad . \quad . \quad . \quad . \quad (7)$$

It is doubtful if a metal surface can be prepared which has an effective conductivity in excess of about 6×10^7 mhos/metre, and hence dielectric materials for which $\tan \delta \sqrt{\epsilon}/(\epsilon - 1)$ is less than 5×10^{-5} at 3000 Mc/s, or 9×10^{-5} at 10000 Mc/s, should prove to be better than metals.

(6) ESTIMATED REDUCTION IN WALL LOSS

By means of dielectric padding, wall loss is reduced by the factor $R = \mathscr{R}Z/\mathscr{R}Z_M$, which, for a single quarter-wave thickness of dielectric in the unbounded case is

$$R = \frac{1}{\epsilon} + \frac{\tan \delta}{\sqrt{\epsilon}} \frac{30\pi^2}{\mathscr{R}Z_M} \quad . \quad . \quad . \quad . \quad . \quad (8)$$

or for an infinity of layers,

$$R = \tan \delta \frac{\sqrt{\epsilon}}{\epsilon - 1} \frac{30\pi^2}{\mathscr{R}Z_M} \quad . \quad . \quad . \quad . \quad . \quad (9)$$

Materials with a low dielectric constant are not of practical interest, since a number of comparatively thick layers would be required. Moreover, if the dielectric constant is near to unity it is necessary for $\tan \delta$ to be exceedingly small for any advantage to be gained. Thus, for example, fused quartz with $\tan \delta$ as low as 6×10^{-5} and a dielectric constant of 3·78 at 3000 Mc/s gives for $\tan \delta \sqrt{\epsilon}/(\epsilon - 1)$ the value 5×10^{-5} and so no improvement is to be expected.

Probably the most suitable materials are ceramics having dielectric constants in excess of 10. A single layer only is required, and the loss reduction factor is given effectively by the second term in (8), namely

$$R = \frac{\tan \delta}{\sqrt{\epsilon}} \frac{30\pi^2}{\mathscr{R}Z_M} \quad . \quad . \quad . \quad . \quad . \quad (10)$$

Ceramics consisting mainly of titanium dioxide are known to have a $\tan \delta$ of between 2×10^{-4} and 3×10^{-4} and an ϵ of

out 90. Such materials may be expected by (10) to give a loss duction of about one-half. An experiment to confirm this is scribed in Section 9.

(7) REFLECTION IN WAVEGUIDE

As has already been stated, eqn. (3) gives the impedance at the rface of a quarter-wave layer of dielectric placed at a distance a quarter of a guide wavelength from a plane end wall in a veguide of any section. In the case of a rectangular guide, the pression also applies to dielectric padding attached to the side lls. Dielectric padding can also be applied to the side wall in guide of circular section, but the analysis of reflections, though alitatively similar, requires a modified treatment especially if dielectric thickness is an appreciable fraction of the guide lius.

For the case of the padded end wall (normal to the guide axis) values of Z_A and Z_D depend on whether an E- or an H-mode being propagated. For an E-mode it is well known that

$$Z_D = \frac{1}{\epsilon}\sqrt{\left\{\frac{\mu_0}{\epsilon_0}[\epsilon - (f_c/f)^2]\right\}}$$

d

$$Z_A = \sqrt{\left\{\frac{\mu_0}{\epsilon_0}[1 - (f_c/f)^2]\right\}}$$

d hence by (3)

$$= \frac{1 - \frac{1}{\epsilon}(f_c/f)^2}{1 - (f_c/f)^2}\frac{1}{\epsilon}Z_M + \sqrt{\left[1 - \frac{1}{\epsilon}(f_c/f)^2\right]}\frac{30\pi^2}{\sqrt{\epsilon}}\tan\delta \quad . \quad (11)$$

ar cut-off ($f \to f_c$) the loss at the metal wall is greater than the unbounded case, but for $\epsilon > 10$ no other change is ticeable. For an H-mode

$$Z_D = \sqrt{\frac{\mu_0/\epsilon_0}{\epsilon - (f_c/f)^2}} \quad \text{and} \quad Z_A = \sqrt{\frac{\mu_0/\epsilon_0}{1 - (f_c/f)^2}}$$

d hence

$$= \frac{1 - (f_c/f)^2}{1 - \frac{1}{\epsilon}(f_c/f)^2}\frac{1}{\epsilon}Z_M + \frac{1}{\sqrt{\left[1 - \frac{1}{\epsilon}(f_c/f)^2\right]}}\frac{30\pi^2}{\sqrt{\epsilon}}\tan\delta \quad . \quad (12)$$

ar cut-off the loss at the metal wall is less than for the bounded case, but again for $\epsilon > 10$ there is little other difence. One may conclude, therefore, that in cases of practical erest the reduction in wall loss to be gained by the use of a rticular dielectric material in a waveguide or resonant cavity little different from that which would be obtained for the case an unbounded wave given by eqn. (10).

APPLICATION OF SPECIALLY PREPARED DIELECTRICS

t is known that the loss factor of a material can be effectively ninished by introducing loss-free air spaces throughout its ume. This procedure must inevitably lower the apparent lectric constant, but, as will now be shown, an overall reducn in the factor $\tan\delta/\sqrt{\epsilon}$ can be obtained.

One simple way in which air spaces can be introduced into a terial is by laminating it. A composite medium formed in this y has been discussed by Harvie,* who has shown that the ective dielectric constant and loss factor depend on the ection of the electric field, the composite medium being mogeneous.

f $k = g/(g + h)$, where g is the thickness of each dielectric layer and h the width of the air space between layers, it can be shown that for an electric field normal to the laminations the effective dielectric constant ϵ_n and loss factor $(\tan\delta)_n$ are given by

$$\epsilon_n = \frac{1}{1 - k + k/\epsilon} \quad . \quad . \quad . \quad . \quad . \quad (13)$$

and

$$(\tan\delta)_n = \tan\delta \, . \, k \, . \, (\epsilon_n/\epsilon) \quad . \quad . \quad . \quad (14)$$

where $\tan\delta$ and ϵ refer to the dielectric forming the laminations. For an electric field parallel to the laminations one obtains

$$\epsilon_T = 1 - k + k\epsilon \quad . \quad . \quad . \quad . \quad . \quad (15)$$

$$(\tan\delta)_T = \tan\delta \, . \, k \, . \, (\epsilon/\epsilon_T) \quad . \quad . \quad . \quad . \quad (16)$$

These quantities were derived by Harvie* on the assumption that the loss factor is small and the thickness of the laminations negligible compared with the wavelength. (In his paper the Q-factor of the material is referred to, which is equal to $1/\tan\delta$.)

For the normal field it follows that

$$\frac{(\tan\delta)_n}{\sqrt{\epsilon_n}} = \frac{\tan\delta}{\sqrt{\epsilon}}k\sqrt{(\epsilon_n/\epsilon)} \quad . \quad . \quad . \quad . \quad (17)$$

The factor $k\sqrt{(\epsilon_n/\epsilon)}$ is always less than unity and hence an improvement in reflectivity must result. For the case of titania, taking ϵ as 90 and ϵ_n as 10, it has the approximate value of one-third. By this means it should be possible to prepare a surface six times as reflective as a metal, a factor of two being obtainable with solid titania.

For an electric field parallel to the laminations,

$$\frac{(\tan\delta)_T}{\sqrt{\epsilon_T}} = \frac{\tan\delta}{\sqrt{\epsilon}}k(\epsilon/\epsilon_T)^{3/2} \quad . \quad . \quad . \quad . \quad (18)$$

No improvement is to be obtained in this case since the factor $k(\epsilon/\epsilon_T)^{3/2}$ is greater than unity in all cases of practical interest.

The reason why an improvement is obtained in the case of the normal field is that the energy density is higher in the air spaces than in the dielectric, and hence the best possible use is made of the air spaces. When the field is tangential the energy density in the dielectric is greater than in the air space and consequently losses are greater. By considering boundary conditions it can be seen that these are, in fact, the limiting cases. In the former the field strength in the dielectric is $1/\epsilon$ times the field strength in the air space, whereas in the latter the field strengths in the air and dielectric are the same. In any other method of aerating the dielectric, the field strength in the dielectric must lie between these limits.

There are obvious practical difficulties in preparing laminated materials suitable for 'padding'. If the dielectric laminations are of a material with a high value of ϵ it can be seen from (13) that ϵ_n is almost independent of ϵ and is given, approximately, by g/h; hence the air space is only $1/\epsilon_n$ times the thickness of a dielectric layer.

For maximum effect the thickness of each dielectric layer requires to be a small fraction, say one-tenth, of the wavelength in the dielectric, and hence the air space is approximately $1/(10\epsilon_n\epsilon^{1/2})$ times the free-space wavelength. For the case of titania at 10 cm, this gives an air space of only about 0·011 cm. The dimension of the air space is extremely important and the task of preparing such laminations is formidable. This places an upper limit to the values of ϵ_n and ϵ which can be used.

(9) EXPERIMENTAL MEASURE OF REFLECTIVITY

It was decided to use a cavity method to measure the reflectivity of a dielectric-padded surface, since free-wave and waveguide methods tend to become difficult and inaccurate when compari-

* R. SHERSBY-HARVIE, R. B., MULLETT, L. B., WALKINSHAW, W., BELL, J. S., and ...CH, B. G.: 'A Theoretical and Experimental Investigation of Anisotropic-Dielectric-... Linear Accelerators', *Proceedings I.E.E.*, Paper No. 2127 M, July, 1956 (104 B,

sons are to be made of highly reflecting surfaces. A cylindrical cavity of circular section and plane end walls was chosen, and the Q-factor of this cavity (Q_M) was measured by a conventional transmission method, coupling loops having been inserted at diametrically opposite positions about half-way along the curved wall. One end wall was then removed and replaced by a dielectric disc of quarter-wave thickness. The curved wall of the cavity was extended for a quarter-wave distance behind the disc and a flat metal plate provided the terminating wall, as shown in Fig. 2. Using the same coupling loop the Q-factor was again determined (Q_D), and from a knowledge of these two quantities it was possible to calculate the change in reflectivity.

Ceramic discs of titanium dioxide (with some stabilizing additives) were available which had a dielectric constant of 93, a $\tan\delta$ of about 3×10^{-4} and a diameter of 3·031 in. The diameter of the cavity was chosen to accommodate this disc, and in the frequency range 3 000 Mc/s $\pm$ 500 Mc/s which could be covered by the available measuring gear there were only two modes which could be used, namely the H_{111} and the E_{011}.

In the H-mode the wall loss is considerably greater on the curved wall than on the end wall. As may be shown by calculation, for a cavity of diameter 3·031 in and for the case when all walls are metal, the loss on the curved wall is about 6·24 times that on one end wall, and consequently the introduction of dielectric padding on one end wall could have only a small effect on the Q-factor of the cavity. For the E-mode the loss at the curved wall is only 2·06 times that at one end wall and hence this mode was chosen.

In order to obtain an accurate comparison between Q_M and Q_D, the same cavity and coupling system was used for each. To aid in assembly a copper band was shrunk on to the ceramic disc and a recess cut at one end of the tube, forming the curved wall of the cavity, to accommodate the copper band, as may be seen in Fig. 2. In experiments to measure Q_M a metal disc, recessed to accommodate the projection of the copper band beyond the

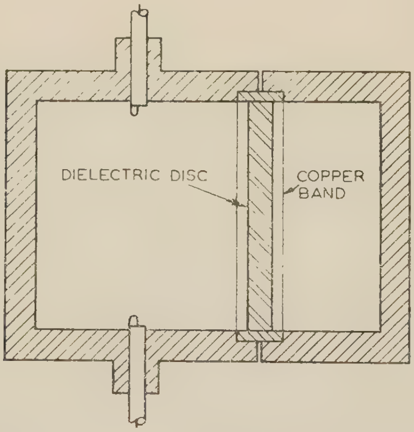


Fig. 2.—Cavity used for reflection measurements.

dielectric surface, was pressed on to form an end wall. The other end wall, which was also detachable, consisted of a flat plate butting against the remote end of the tube. Pressure plates were then applied at each end of the cavity and the whole was drawn together with tie rods.

This method has been used successfully in many other cavity experiments using titania discs. Reliable, repeatable results can easily be obtained, but care must be taken to ensure that all surfaces are clean and accurate in shape.

Measurements of Q_M and Q_D were taken over a range of frequencies, the resonant frequency of the cavity being altered by

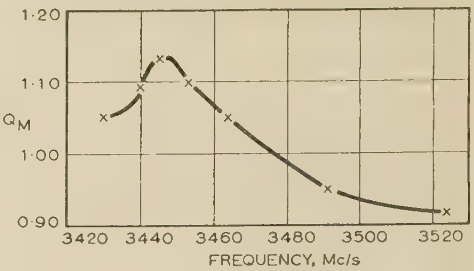


Fig. 3.—Ratio of cavity Q-factors as a function of frequency.

Q_D refers to cavity with dielectric padding.
Q_M refers to the all-metal cavity.

turning off metal from one end in order to reduce its length. The results are shown in Fig. 3. As would be expected, a marked improvement is to be found over a narrow band of frequencies, the maximum value of the ratio Q_D/Q_M being about 1·13.

Since the volume, frequency and mode pattern of the cavity are the same in measuring both Q_D and Q_M, the ratio of the Q's is inversely proportional to that of the wall loss in the two cases. Thus, if x represents the factor by which the loss in one end wall is diminished by the dielectric padding,

$$\frac{Q_D}{Q_M} = \frac{2\cdot06 + 2}{2\cdot06 + 1 + x}$$

so that for $Q_D/Q_M = 1\cdot13$, $x = 0\cdot53$ and the padded wall is approximately twice as reflective as the metal.

(10) CONCLUSION

The purpose of this work has been to examine the theory of dielectric padding and to demonstrate that it is practicable. It is unlikely that there are applications at frequencies below about 3 000 Mc/s, and even at that frequency the improvement in reflection by a factor of two which was obtained with titania is exceptional and no material was found which could exceed this figure. At higher frequencies, owing to the deterioration in the reflection from a metal, the position appears to be much more hopeful, but as yet no experimental measurements have been made.

The quality of a dielectric material for use in padding depends almost entirely on the factor $\tan\delta/\sqrt{\epsilon}$, and it may be possible to manufacture materials which are better than existing ones in this respect. The fact that a considerable improvement, in theory at least, may be obtained by laminating any particular material provides support for this expectation.

It has been shown that dielectric padding gives a maximum advantage at a fixed frequency and provides a useful advantage only over a narrow frequency band. This is inherent in the method, and any attempt to improve the bandwidth, for example by using multiple slightly detuned layers, must inevitably result in a loss in reflectivity. For devices working at a fixed frequency, the tolerances on the dimensions of the padding are normally less exacting than for a simple metal wall. The most important dimension is the thickness of the dielectric slab, but this does not present any special difficulty when ceramic materials are used, since they can easily be ground to exceedingly fine limits.

(11) ACKNOWLEDGMENTS

The work described was carried out at Queen Mary College and was supported by an A.E.R.E. contract. The titania ceramic used was specially prepared by Dr. E. Schaefer, with whom the authors had many helpful discussions for which they wish to express their thanks.

21.317.794:621.372.822

The Institution of Electrical Engineers
Paper No. 2488 R
Jan. 1958

TRANSVERSE FILM BOLOMETERS FOR THE MEASUREMENT OF POWER IN RECTANGULAR WAVEGUIDES

By J. A. LANE, M.Sc., Associate Member.

(*The paper was first received* 25*th September, and in revised form* 30*th October*, 1957.)

SUMMARY

The paper describes a simple standard technique, using thin metallic ms sputtered on mica, for the direct measurement of powers in the nge 1–100 mW. The experiments were carried out at a frequency of 0 Gc/s (wavelength 3 cm), but the methods used are of general applica-on in rectangular waveguides.

The absorbing films consisted of a platinum deposit, of the order 10^{-6} cm thick, on a mica strip 0·3–0·4 cm wide, located sym-etrically in the transverse plane. The change in resistance or rise temperature can be used as a measure of input power, which is etermined by a d.c. calibration. The operating bandwidth, for a ven waveguide, is greater than that normally available with thermistors ad bolometers, and the techniques described seem especially con-nient for the calibration of these instruments.

Good agreement has been obtained in a comparison with calorimeters ad a force-operated wattmeter, and the results indicate that an error mi of not more than ±3% can be achieved at a frequency of 10 Gc/s.

(1) INTRODUCTION

At frequencies above a few hundred megacycles per second e efficiency of transmitting and receiving equipment is generally tablished by measurements of power, since in this region of the dio-frequency spectrum the basic quantities of current and bltage are difficult to determine. As a result, much effort has ready been devoted to the development of standard techniques power measurement, especially for frequencies above 2 Gc/s vavelengths below 15 cm). The constant-flow water calori-eter, for example, is frequently used for this purpose in wave-iide transmission lines. In addition, torque-operated watt-eters of the single- or double-vane[1,2] type are now available the United Kingdom for frequencies near 10 Gc/s, and are, in any ways, more convenient in use than calorimeters. All these struments, however, require powers of several watts or more r their operation and many measurements in microwave evelopment are carried out at much lower levels. There is a ed, therefore, for power-measuring standards at all powers low about 1 watt, and especially for powers of a few milliwatts. ome work has already been done with this aim in view; for ample, a microwave microcalorimeter has been developed at e National Bureau of Standards for use[3] at 10 Gc/s, and odified force-operated methods for use at relatively low levels ve been investigated by Cullen and French,[4] and by Bailey.[5] evertheless, a simple method suitable for general use is still important requirement.

The paper describes a technique, using thin metallic films, ich has been developed in an attempt to meet this need.[6] It as considered that a method having an error limit of 2–3% or s would represent a useful advance in measuring technique, pecially if this performance could be achieved at frequencies the order of 10 Gc/s and above.

Wr tten contributions on papers published without being read at meetings are ite for consideration with a view to publication.

h paper is an official communication from the Radio Research Station, Depart-nt of Scientific and Industrial Research.

(2) BASIC CONSIDERATIONS

The method generally used at the present time for the measurement of power at the milliwatt level makes use of thermistor beads or bolometer wires calibrated by d.c. power. These resistance-type milliwattmeters have, however, several limitations as standard instruments at frequencies above 3 Gc/s. They can, of course, be calibrated against the high-level standards by using calibrated attenuators, but the experimental difficulties encountered are considerable.

In considering possible alternative techniques, the advantages of a film bolometer, with a thickness appreciably less than the skin depth, were realized at an early stage in the present investigation. Little work appears to have been done on film bolometers for waveguide applications, and the only published information known to the author relates to extended films, similar to those used in variable attenuators, placed parallel to the waveguide axis. This arrangement was used by Collard in an instrument termed the enthrakometer, a fraction of the input power being absorbed in a film on the side wall of the waveguide.[7] Similarly, resistive films have also been used in an air thermometer for the measurement of powers of 10–100 mW at a frequency[8] of 10 Gc/s. The main difference between these two instruments on the one hand and the instrument described in the present paper is in the arrangement of the film and in the methods used for measuring the power absorbed.

It is well known that a thin film in the transverse cross-section of a waveguide, when followed by a perfectly reflecting plunger at a distance of $(2n-1)\lambda_g/4$, where $n = 1, 2, 3$, etc., and λ_g is the wavelength in the guide, will act as a non-reflecting termination, provided that the film has a uniform surface resistivity equal to the wave impedance, Z, of the mode being transmitted. For the H_{01} mode in rectangular waveguide we have

$$Z_{H_{01}} = \frac{E_{transverse}}{H_{transverse}} = 377\lambda_g/\lambda \text{ ohms} \quad . \quad . \quad . \quad (1)$$

As pointed out by Collard,[7] an arrangement of this kind should facilitate the measurement of relatively low powers. In such a film, however, the axial component of the Poynting vector will vary in magnitude over the transverse cross-section, the power dissipation per unit area being a maximum along a central line parallel to the narrow wall in the case of the dominant mode. It was considered, therefore, that the requirement of maximum sensitivity, coupled with the ideal of pure substitution of d.c. power for microwave power, would best be achieved by using a relatively narrow strip in the transverse plane, located symmetrically in a region in which the transverse electric and magnetic fields are very nearly uniform. A waveguide obstacle of this nature has not been investigated either theoretically or experimentally, so far as the author is aware. In fact, it is not immediately obvious that a non-reflecting termination can be obtained in this way. However, as will be evident below, the input-voltage standing-wave ratio (E_{min}/E_{max}) can be made comparable with the values normally achieved in existing milliwattmeters. Moreover, this performance can be obtained without the need for impedance transformers.

Since the temperature gradients over the transverse film are very nearly identical under d.c. and high-frequency conditions, it is merely a matter of convenience whether the change in resistance or the temperature rise at a point is used as an indication of the power absorbed. Both methods have been used in the present investigation, although the latter technique is preferred since a direct indication of the power can be obtained by using a thermo-junction and galvanometer. This procedure is similar to that followed, for example, by Guild in the absolute temperature-drift radiometer developed for use at optical frequencies.[9]

(3) DETAILS OF DESIGN AND CALIBRATION

All the films used in this work were prepared in the Light Division of the National Physical Laboratory and consisted of a platinum deposit (of the order of 10^{-6} cm thick) on mica strips 10^{-2} cm thick. Film widths of 0·3–0·4 cm were found to be convenient in standard 3 cm band waveguide. The platinum was deposited by a process of sputtering in a vacuum chamber at a pressure of the order of 10^{-2} mm Hg, the resistivity being controlled in the first instance by regulation of the sputtering time. Subsequent adjustments of resistivity to obtain an optimum value were made by heating the film on a hot-plate. Gold deposits were prepared on the ends of the film to which d.c. contacts were attached in the final assembly. Typical dimensions for the completed film are shown in Fig. 1.

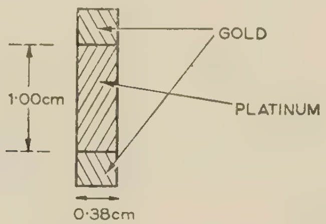


Fig. 1.—Dimensions of typical film bolometer.

A preliminary investigation was made of the effect of variations in film width and resistivity on the measured normalized impedance of the film when located in the transverse plane. It is only necessary here to summarize the results obtained for the strips shown in Fig. 1, and typical values obtained at a frequency of 9·2 Gc/s (wavelength 3·26 cm) are given in Table 1.

Table 1

MEASURED NORMALIZED IMPEDANCE OF BOLOMETER FILM: $f = 9{\cdot}2$ Gc/s

D.C. resistance	Measured normalized impedance
ohms	
260	$0{\cdot}52 + j0{\cdot}2$
370	$0{\cdot}70 + j0{\cdot}2$
580	$1{\cdot}13 + j0{\cdot}2$

The film thus behaves as a resistance in series with an inductance, the value of the latter quantity being approximately constant for a film of given width. A film having a d.c. resistance of 480–500 ohms should therefore provide a non-reflecting waveguide load at a frequency of 9·2 Gc/s when followed by a reflecting plunger at a distance slightly greater than $\lambda_g/4$. The calculation of the input impedance of the complete unit is summarized in the Appendix (Section 8), and the results are shown in Fig. 2. The curves were calculated for a film of

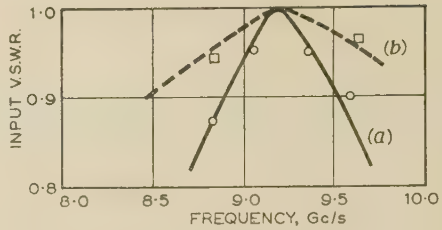


Fig. 2.—Variation of input v.s.w.r. with frequency for bolometer of optimum resistance.

(*a*) For fixed plunger.
(*b*) For movable plunger.
○□ Experimental points.

optimum resistivity; curve (*a*) shows the variation of input-voltage standing-wave ratio (v.s.w.r.) for a fixed position of plunger, and curve (*b*) is the corresponding variation with the plunger adjusted for minimum reflection at every frequency. Even with a fixed plunger, a bandwidth of approximately 600 Mc/s is obtained for an input v.s.w.r. greater than 0·9. In practice, owing to small departures in resistivity from the optimum value and other imperfections in the mount assembly, the bandwidth is somewhat less than that indicated by Fig. 2. Typical results for a complete instrument are given in the next Section.

The complete film was clamped in a shallow recess between two rectangular flanges with the platinum and gold deposits facing the input flange. One end was insulated from the waveguide by a strip of mica, 4×10^{-3} cm thick, and a d.c. connection made to a thin strip of copper foil extending outside the flanges. The final adjustments in the location of the film were made using a low-power microscope, and a single copper-Eureka thermo-junction constructed from 40 s.w.g. wire was attached to the centre of the platinum deposit by a minute quantity of suitable adhesive. In order to reduce reflections from the thermo-junction wires to a minimum, the latter were arranged to be parallel to the broad face of the waveguide, i.e. perpendicular to the transverse electric field. In observing the steady-state temperature rise, some temperature compensation was obtained by the use of a similar thermo-junction attached to, but insulated from, the outside of the waveguide at a point as close as possible to the film. The whole unit was enclosed in a small container packed with cotton-wool lagging. Polished copper was used for the waveguide and plunger, the latter being fixed at a distance of 1·3 cm behind the film.

Platinum films mounted in this way have been used, as mentioned above, both as resistance-type milliwattmeters and as direct-reading instruments. In the latter method, powers of 1–100 mW can be read directly from the same galvanometer scale by using suitable resistances in series with the thermo-junction. As a result of overheating, there is some danger of damage to the film at power levels much above 100 mW in the present technique, although powers of 200–300 mW have been dissipated in some films for a brief period. Typical calibration results for two conditions of sensitivity are given in Fig. 3. Using a sensitive reflecting-mirror galvanometer it was possible to obtain a scale deflection of approximately 10 cm/mW under conditions of maximum sensitivity. The effective time-constant of the system varied from 15 sec under these conditions to about 3 sec when powers of 10 mW or more were being measured. This relatively sluggish response is probably not a serious limitation in laboratory measurements of the kind for which the instrument is intended. (Subsequent experiments have shown that a smaller time-constant is obtained, with an increased sensitivity, by using a d.c. amplifier in the indicating circuit.)

The d.c. calibration, having an estimated error limit of 0·4–0·5%, was usually carried out at each series of measure-

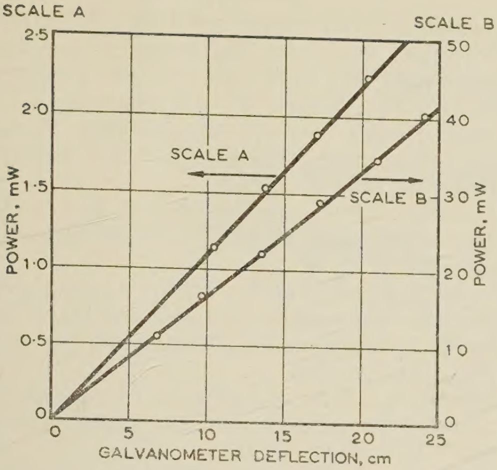


Fig. 3.—Calibration of typical film bolometer.

nts. Although no evidence is yet available on the degree of ability of film resistance over a long period, no change has en detected in the properties of a film over a period of one or o months.

When used as a resistance-type milliwattmeter, the film was nnected into a simple d.c. Wheatstone bridge having a variable cade resistance box in series with the battery. The input wer was determined by observing the change necessary in this sistance to maintain the bridge in a state of balance. Powers 10–100 mW could be measured in this way, but varying bient temperature proved troublesome at much lower levels. bsidiary experiments included determinations of the input .w.r. for the complete instrument, and measurements of the idual mount loss due to power dissipation in the waveguide nges, short-circuiting plunger, etc. Fig. 4 illustrates the input

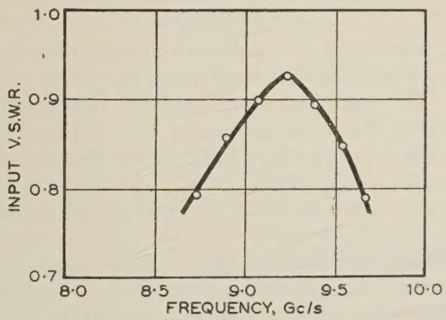


Fig. 4.—Variation of input v.s.w.r. with frequency for typical instrument.

w.r. obtained with a typical instrument and may be compared h the optimum values given in Fig. 2. From measurements the input impedance in the absence of the film, the mount loss s estimated to be not more than 0·3%.

(4) COMPARISON WITH EXISTING EQUIPMENT

Experience has shown that power measurements at microwave frequencies are particularly liable to concealed systematic errors which are often revealed only after much careful investigation. Comparisons between as many methods as possible are of great importance in assessing the accuracy of a new technique, especially if the instruments being compared differ in their principle of operation. Simultaneous measurements have therefore been made using film bolometers of the type described and wire bolometers, water calorimeters, and a force-operated wattmeter of the double-vane type. Of particular importance are those results obtained with water calorimeters and the force-operated wattmeter, for experience has shown that these two instruments are reference standards of comparable accuracy at a frequency of 10 Gc/s.

In the first series of measurements, comparisons were made between a film bolometer and calorimetric equipment previously developed for the calibration of resistance-type milliwattmeters.[10] Calibrated directional couplers and vane attenuators were used as necessary to establish powers in the range 1–120 mW. The absolute accuracy of this power level varied slightly according to the experimental conditions, but the inaccuracy was estimated to be less than 3% at any power level. Typical results on the performance of the film bolometer are given in Table 2 for what are termed, for the sake of convenience, the direct-reading and resistance methods.

Table 2

COMPARISON OF METHODS OF USE OF FILM BOLOMETER

Method	Film bolometer reading, P_b	Calorimeter reading, P_c	Average difference $(P_b - P_c)\,100/P_c$
	mW	mW	
Direct-reading	1·84	1·88	
	2·65	2·68	
	12·2	12·1	
	30·0	30·4	−0·3
	48·4	48·1	
	76·7	76·2	
	121·1	120·7	
Resistance	12·0	12·2	
	19·2	19·3	
	24·1	24·4	
	38·6	38·5	−0·9
	46·7	47·9	
	95·2	95·7	
	119·9	120·3	

The variations in the individual values of $(P_b - P_c)\,100/P_c$, expressed as a standard deviation from the mean difference, are 1·1 and 0·9 for the two series; there is little significance therefore in the difference between the figures in the final column.

After these measurements, the calorimetric equipment was dismantled and subsequently rebuilt with minor modifications of design. Simultaneous comparisons were then made between a film bolometer, the water calorimeter, and a torque-operated wattmeter of the double-vane type. The complete assembly for

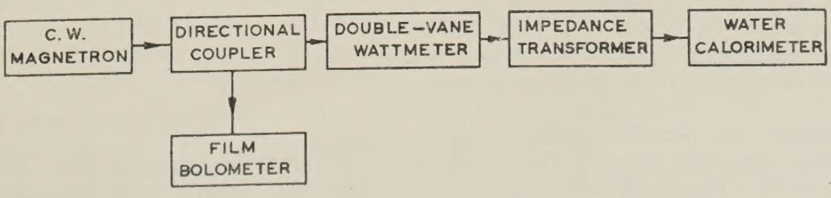


Fig. 5.—Comparison of film bolometer with water calorimeter and double-vane wattmeter.

these experiments is shown in outline in Fig. 5. The experimental procedure was similar to that followed in previous comparisons of this kind,[2,11] and only the final results are summarized here. The figures given in Table 3 specify the net input power to the film bolometer, corrections having been made to allow for power losses in the impedance transformer and in the vane wattmeter. It may be mentioned here that it proved possible to obtain accurate readings with the vane wattmeter at a power level as low as 3 watts by using a fine hair-line in a galvanometer lamp which was appreciably brighter than normal. Typical results of the comparison are given in Table 3; here the film-bolometer results were all obtained by the direct-reading method.

Table 3

COMPARISON OF FILM BOLOMETER WITH WATER CALORIMETER AND VANE WATTMETER

Film bolometer reading, P_b	Calorimeter reading, P_c	Vane wattmeter reading, P_v
mW	mW	mW
1·05	1·02	1·02
1·71	1·76	1·73
2·42	2·53	2·48
19·1	19·4	19·5
20·6	20·9	20·7
29·4	30·0	29·9

The average value of $(P_v - P_c)\,100/P_c$ is $-0{\cdot}8$, and this may be regarded as further experimental confirmation of the accuracy of the torque-operated type of wattmeter. The average difference between the film-bolometer reading, P_b, and the mean reading of the other two instruments is $-1{\cdot}2\%$. The film bolometer therefore agrees with the high-level standards within the limits of accuracy with which the instruments can be compared. The random errors in the comparisons are largely due to the effect of variations in local ambient temperature on the film-bolometer reading, but these errors could almost certainly be reduced by improved methods of temperature compensation.

(5) CONCLUSIONS

The experiments described have established that transverse-film bolometers possess many advantages for the measurement of powers of a few milliwatts in waveguide transmission lines. By using thin films of platinum sputtered on mica, powers in the range 1–100 mW can be measured in terms of a d.c. calibration with an error limit which is probably less than $\pm 3\%$ at 10 Gc/s. No impedance transformers are required for most applications and the technique should be equally reliable for both c.w. and pulsed power in rectangular waveguide at any frequency. By improving the design of the mount assembly it should be possible to increase the accuracy still further. Further development at frequencies above 10 Gc/s would seem well worth while.

(6) ACKNOWLEDGMENTS

The author is indebted to Mr. J. S. Preston of the National Physical Laboratory for preparing the metallic films and for many helpful discussions, and to the Wayne Kerr Laboratories Ltd. for the loan of a double-vane wattmeter. The advice given by Dr. J. A. Saxton in the preparation of the paper is also gratefully acknowledged.

The work described above was carried out as part of the programme of the Radio Research Board. The paper is published by permission of the Director of Radio Research of the Department of Scientific and Industrial Research.

(7) REFERENCES

(1) CULLEN, A. L., and STEPHENSON, I. M.: 'A Torque-Operated Wattmeter for 3 cm Microwaves', *Proceedings I.E.E.*, Monograph No. 42 M, July, 1952 (**99**, Part IV, p. 294).

(2) CULLEN, A. L., ROGAL, B., and OKAMURA, S.: 'A Wide-Band Double-Vane Torque-Operated Wattmeter for 3 cm Microwaves', *Proceedings of the Institute of Radio Engineers* (to be published).

(3) MACPHERSON, A. C., and KERNS, D. M.: 'A Microwave Microcalorimeter', *Review of Scientific Instruments*, 1955, **26**, p. 27.

(4) CULLEN, A. L., and FRENCH, H. A.: 'An Instrument for the Absolute Measurement of Low-Level Power in the 3 cm Band', *Proceedings I.E.E.*, Monograph No. 237 R, May, 1957 (**104 C**, p. 456).

(5) BAILEY, R. A.: 'A Resonant-Cavity Torque-Operated Wattmeter for Microwave Power', *ibid.*, Monograph No. 138 R, June, 1955 (**103 C**, p. 59).

(6) LANE, J. A.: 'A Film Radiometer for Centimetre Wavelengths', *Nature*, 1956, **177**, p. 392.

(7) COLLARD, J.: 'The Enthrakometer, an Instrument for the Measurement of Power in Rectangular Waveguides', *Journal I.E.E.*, 1946, **93**, Part IIIA, p. 1399.

(8) GORDON-SMITH, A. C.: 'A Milliwattmeter for Centimetre Wavelengths', *ibid.*, Paper No. 1888 R, September, 1955 (**102 B**, p. 685).

(9) GUILD, J.: 'Investigations in Absolute Radiometry', *Proceedings of the Royal Society* A, 1937, **161**, No. 904, p. 1.

(10) LANE, J. A.: 'The Measurement of Power at a Wavelength of 3 cm by Thermistors and Bolometers', *Proceedings I.E.E.*, Paper No. 1918 R, November, 1955 (**102 B**, p. 819).

(11) BAILEY, R. A., FRENCH, H. A., and LANE, J. A.: 'The Comparison and Calibration of Power-Measuring Equipment at Wavelengths of 3 cm and 10 cm', *ibid.*, Paper No. 1689 R, September, 1954 (**101**, Part III, p. 325).

(8) APPENDIX

The input impedance of a transverse film in front of a reflecting plunger can be calculated in the following way:

Consider, for example, a normalized shunt impedance of $0{\cdot}95 + j0{\cdot}20$ (a shunt admittance of $1{\cdot}0 - j0{\cdot}22$), corresponding to a film resistance of 480 ohms, at a frequency of 9·2 Gc/s ($\lambda = 3{\cdot}26$ cm). The normalized input admittance at the plane of the film can be made unity by locating the plunger 1·30 cm (i.e. $0{\cdot}28\lambda_g$) behind the film.

For a fixed plunger, two factors will contribute to the frequency variation of input impedance or admittance. The characteristic wave admittance is frequency-dependent, as is the susceptance, at the plane of the film, of the short-circuited line. The value of characteristic wave admittance, Y_1, at a free-space wavelength λ_1, relative to the value Y_0 at a free-space wavelength λ_0, is given by

$$Y_1 = Y_0 \frac{\lambda_1 \lambda_{0g}}{\lambda_0 \lambda_{1g}} \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

where λ_{0g}, λ_{1g} are the wavelengths in the guide. The film admittance can therefore be expressed relative to the characteristic wave admittance at any desired frequency. The variation in the susceptance component, B, due to the short-circuited line can be calculated from the relation

$$B = -j \cot (2\pi l/\lambda_g) \quad . \quad . \quad . \quad . \quad . \quad (3)$$

where l is the length of the line. Alternatively, this variation can be derived directly from the polar form of transmission-line chart.

PROCEEDINGS OF THE INSTITUTION OF ELECTRICAL ENGINEERS

Part B. RADIO AND ELECTRONIC ENGINEERING (INCLUDING COMMUNICATION ENGINEERING), JANUARY 1958

CONTENTS

Bibliographical References.—It is requested that bibliographical reference to an Institution paper should always include the serial number of the paper and the month and year of publication, which will be found at the top right-hand corner of the first page of the paper. This information should precede the reference to the Volume and Part.

Example.—SMITH, J.: 'Reflections from the Ionosphere', *Proceedings I.E.E.*, Paper No. 3001 R, December, 1954 (102 B, p. 1234).

Published by The Institution, Savoy Place, London, W.C.2. *Telephone:* Temple Bar 7676. *Telegrams:* "Voltampere, Phone, London."
Printed by Unwin Brothers Limited, Woking and London.